Author:
Daw, Katrina

Title:
The development of semantic cognition

Effects of typicality on semantically-driven tasks in children aged five to ten

The development of semantic cognition: Effects of typicality on semantically-driven tasks in children aged five to ten

By: Katrina Jayne Daw

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Master of Science in the Faculty of Life Sciences, School of Psychological Science. Submitted: 12/2018 Word count: 26 765

Abstract

This study investigated experiential and controlled cognition accounts of conceptual knowledge development in a sample of 107 children aged five to ten. Four semantically-driven tasks: naming, item matching, sorting and category matching explored varying typicality of a common set of items, whilst controlling for familiarity. A further object-use task explored matching of a probe to a functionally-related target from amongst semantically-related distractors. Variable performance on tasks using the same items implicated task demands, not just variation in vocabulary knowledge. Naming accuracy showed a graded typicality advantage. Partial support was found for recruitment of controlled cognition in the remaining tasks. For example, in item matching, alongside proximal distractors, there was a greater cost to accuracy for more typical probes. For sorting, accuracy was more variable when categories were more specific, indicating a requirement for more controlled cognition. Matching by specific sortal improved with greater distinctiveness of the most and least typical items for all ages, despite their familiarity. For category and object matching, proximal distractors impeded performance, due to a greater requirement for controlled cognition. Contrary to item matching, the effects of proximal distractors on category matching were greater for less typical items where category membership was less direct, suggesting a role for controlled cognition.

The findings further suggest a tentative conclusion that younger children (60-78 months) were less able to recruit controlled cognition. When sorting, these children benefited from a specific sortal term, where the relationship between the item and the category is more proximal than for general sorting, offsetting the requirement for controlled cognition. Older children were also better at matching by category and function; tasks with greater control requirements due to more distant relations between probe and target. The findings are discussed in relation to the controlled semantic cognition (CSC) framework (Rogers, Patterson, Jefferies & Lambon Ralph, 2015).

Dedication and Acknowledgements

Thank you to my main supervisor: Dr Josie Briscoe for all of her support and feedback. I have needed to learn many new skills in order to carry out this research and I am very grateful to her for sharing her expertise with me. I have also received assistance with RStudio from Miss Sarah Von Grebmer Zu Wolfsthurn. She has taught me how to do different things in RStudio and this has helped me to overcome the problems I was having with data analysis.

This research would not have been possible without the participation of the following schools: Bourton on the Water Primary Academy, Churchdown Village Infant School, Churchdown Village Junior School, St. James' Church of England Primary School, The Richard Pate School and Warden Hill Primary School. We really appreciate the time that staff at these schools have given up to allow us to visit their schools and test their children. It was extremely kind of Mr Sam Williams and Miss Charlie Rand to let me test not one but two of their classes. I would also like to thank Mr Adams, Mrs Karen Priest, Mrs Sue New, Mrs Anya Wood and Mrs Sue Wade for their assistance with recruiting pupils and schools.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .... DATE: ....

Table of Contents

List of tables and figures

1 Introduction

1.1. The challenges of early and late conceptual development

One of the greatest challenges of childhood is acquiring and using conceptual knowledge of, and about, the world (Rosch et al., 1976). The world contains an enormous variety of objects that can be represented as concepts: mental representations of the properties of objects (Hoffman, McClelland & Lambon Ralph, 2018). By using naturally-occurring structures in objects, including their locations and functions in relation to other objects, it is possible for a naïve learner to group items into categories (Rosch, Mervis, Gray, Johnson & Boyes-Braem, 1976). The similarity of bundles of features can be used to form meaningful associations between objects and comprise a basis for categories. For example, features such as 'feathers', 'wings' and 'beaks' often co-occur together and entities with all three features can form the 'bird' category. Traditionally, theories of early conceptual development have focused on the role of experience to explain how children acquire and use concepts and categories. In recent years, executive cognitive abilities have been implicated in the ability to organise and use concepts efficiently (Jefferies & Lambon Ralph, 2006). This thesis will focus on how children use their conceptual knowledge in an efficient and goal-sensitive way, to describe how developmental change might occur in the broader context of conceptual development.

1.2 Early categorisation and object naming

Young infants are aware of the perceptual differences between certain animals by 3 months of age (Eimas & Quinn, 1994). When they are familiarised with horses, they prefer to look at cats, zebras and giraffes, rather than horses, on test trials, showing sensitivity to animate features. By 9 months of age, infants are aware of conceptual, as well as perceptual differences. They know that animals and vehicles not only look different but are distinct concepts (Mandler & McDonough, 1993). In this experiment infants were familiarised with four exemplar objects from either an animal or a vehicle category. After this, they were presented with an object from the same category followed by an object from the contrasting category. Infants spent longer examining the item from the contrasting category, indicating that they could discriminate animals from vehicles. This preference was evident even when the animals and vehicles looked similar. The same experiment showed that 9 and 11-month-olds did not discriminate dogs from fish or rabbits. These findings arguably suggest that global categories (such as animals and vehicles) are developed earlier than basic categories that represent the most frequently occurring category members (see also Quinn, Westerlund & Nelson, 2006). Whilst the developmental trajectory of category development is clear i.e. that general categories are established before basic categories, some researchers argue that these first categories are perceptually (rather than conceptually) grounded. This raises further questions about the nature of early concepts during infancy and early childhood.

An important cue to conceptual knowledge is children's emerging ability to identify a concept by name. Early categorisation is related to naming ability in the second year (Gopnik & Meltzoff, 1992). Children aged approximately 16 to 18 months were more successful at categorisation tasks, that required grouping identical items together (e.g. Raggedy Andy figures and small toy trains), if they had larger vocabularies. Gopnik and Meltzoff found that naming ability was related to categorisation, even when items differed on other dimensions such as colour. Other evidence suggests that by 20 months of age infants can use labels provided by adults to categorise objects (Nazzi & Gopnik, 2001). In this experiment, the 20-month-olds were able to match items based on naming information, even though they were perceptually dissimilar. It is not only labelling that is important, infants of a similar age (18 months) are aware of what adults are attending to (Baldwin, 1993). Baldwin (1993) found that 18-month-olds were able to work out which toy a label referred to. This was evident even when the infant and the experimenter were looking at different objects at the point of labelling. Of course, providing names for objects can guide learning by helping infants to make distinctions between different concepts. It can also aid their discovery of features that are present in different items. For example, when 17-month-olds were allowed to play with, and exposed to the names of, objects that were similar in shape but differed by colour and texture, they could extend names to novel items that were also similar in shape (Smith, Jones, Landau, Gershkoff-Stowe & Samuelson, 2002). It seemed that labelling the objects and giving the children practical experience with them was sufficient to orient infants to specific attributes, such as shape. The evidence suggests that experience can lead to the discovery of new word-object mappings in unambiguous situations. However, what happens in the ambiguous setting of the real world? When a parent labels an object, it is not always obvious what they are referring to. However, young children frequently work out what words mean from limited information (Deák, 2000). This suggests that factors other than experience are involved in the learning process.

1.3 Concepts, categories and function in preschool children

Objects can also be divided into natural and artefactual kinds. Kinds that exist in nature are referred to as natural kinds, whereas artefactual kinds are constructed (Neary, Van de Vondervoort & Friedman, 2012). By 3-4 years of age, children are aware that artefacts, unlike natural kinds, are made by humans (Gelman, 1988). It has been found that young children show an advantage for natural kinds in induction tasks (Freeman & Sera, 1996). In this study, 3-5-year-olds were asked questions about machines and animals that the experimenter had not given them information about. The questions either related to properties of animals or machines. For example, an animal question referred to whether the object 'had a stomach inside'. A machine question concerned whether the item 'would rust if it were left out in the rain'. Three-year-olds achieved higher levels of accuracy for questions that related to properties of animals, as compared to machines. By 18 months of age infants can distinguish between animates and inanimates (Rostad, Yott & Poulin-Dubois, 2012). After this distinction is made, an advantage for animates emerges (Funnell, Hughes & Woodcock, 2006). Children between the ages of 3 years 7 months and 4 years 6 months are more accurate at naming animals than implements or vehicles. The authors argued that the superior performance with animals was a result of exposure to them through looking at picture books and going to the zoo. They emphasised the significance of experience in conceptual acquisition, rather than intrinsic differences between kinds.

Experience is also significant for learning about the functions of objects. For many artefacts, it has been noted that there is a causal relationship between an object's shape and its function (Ware & Booth, 2010). This suggests that attending to object shape can facilitate the acquisition of object functions. By 3-4 years of age, preschool children showed awareness that perceptual features of objects relate to their functions, when asked directly about an object (Asher & Kemler Nelson, 2008). Objects were either assigned functions that accounted for their perceptual or structural features, or the child was provided with a function that ignored the structural features of the object. In the latter condition, children were more likely to ask further questions. This shows that they were not satisfied with the answer they had been given and they appeared to want an explanation for the parts that the first response had ignored. In other words, 4-5-year-old children were biased towards the intended functions of objects and did not easily accept current, contradictory functions. The fact that children show a desire to resolve mismatches between perceptual features and function suggests that function can be an important component of concepts. Further support for this arises from a categorisation study of natural and artefactual kinds (Pauen, 1996). If perceptually dissimilar artefacts shared a function, then children did not accept separating these objects into different groups. This was not true for natural kinds; children were happy for these to be separated if they looked different. Whereas function is a salient property when categorising artefacts, structural or superficial features are important for classifying natural kinds. In a related study with 17-month-olds (Ware & Booth, 2010), children's observation of shape-based object functions in action and their active experience of objects helped them to make matches according to shape. This evidence suggests that knowledge of functional properties can lead to the realisation that items can be categorised according to shape. Categorisation of this kind is based on conceptual ideas, such as what objects are used for, rather than simply perceptual similarity.

1.4 Concepts as learned by association

Objects rarely, if ever, occur in isolation. Therefore, conceptual knowledge must also arise from early exposure to objects in situ. The co-occurrence of objects can provide children with clues as to what unfamiliar objects are; links can be made between items that are found in the same location, are used in the same way or for the same purpose. One view is that once these associative links have been made they can facilitate the discrimination and matching of objects. Young children demonstrate a clear preference for thematic knowledge of objects, for example, when matching probes to objects (Fenson, Vella & Kennedy, 1989). Even children as young as 2 years can match probes to targets that are thematically related, such as sunglasses and eyes. With a thematic relation, young children were successful 60% of the time. However, their aptitude for thematic matching was not repeated when matching by category; their accuracy fell to 28% when the target was perceptually dissimilar to the probe from the same category. The preference for thematic matching persists with older children (Smiley & Brown, 1979). A free recall investigation also provided evidence that young children like to group items based on thematic relationships (Denney & Ziobrowski, 1972). Four- and five-year-olds were more likely than adults to recall thematically related words consecutively.

A more contemporary view is that associative learning through observation also allows generation of the properties of objects, therefore generating access to object identity. A recent model by Rogers and McClelland (2008) describes this process. The theory is a variation of one put forward by Rumelhart (1990, as cited in Rogers & McClelland, 2008). His network is made up of units, which are in layers. The units are connected. The labels given to the layers are: item, relation, representation, hidden and attribute. Attribute units are in the output. The network can be probed by presenting an item and a corresponding relation. For example, the item may be *canary* and the relation may be *can*. *Canary* is represented in one of the item units and *can* is represented in one of the relation units. These units will be activated when *canary* and *can* are presented. The activation spreads forward through the network and certain attribute units should be turned on. For example, *fly* should be turned on because canaries can fly. This will only be true if the individual has learned that canaries can actually fly. They may learn this through interpreting the experience of seeing a canary fly, or they may be exposed to the conceptual notion that canaries can fly. Activation within the model is altered by connection weights, which change with experience. According to this view, any developmental changes in children's learning of objects is really an emerging feature of learning about the co-occurrence of object properties.

Rogers and McClelland (2008) proposed that infants and young children learn properties that are shared by semantically related items more quickly because of the effect of experience on connection weights. For example, when a child sees a canary flying, this experience will increase the connection weights of the item-relation-attribute. When a child learns something about an item belonging to a category, this allows the child to generalise this piece of information to all members of that category. For example, when they see a canary pecking, they can correctly generalise this to all birds and 'learn' that all birds can peck. The network inside the child has been taught this through the process of generalisation and activation. When a child sees other birds pecking, this strengthens the connection weights of different bird items, the relation *can* and the attribute *peck*. When a child learns a property that individuates canaries from other birds, such as the colour of canaries, the child will incorrectly generalise this to all birds. The child will think that all birds are yellow. Upon seeing a bird of a different colour, the connection weights involved in *bird-yellow* will be weakened. For example, if the child sees a penguin, the connection weights will be affected. They will have been strengthened by seeing a yellow canary, but this effect will be reversed when the black and white penguin is seen. The child learns the property that all birds share, pecking, faster than the property (of colour) that differentiates canaries from other birds. This is because every time the child sees a different bird pecking, the relevant connection weights are strengthened.

An important element of this model is that it highlights the role of experience in learning about relationships between objects. Once a fact is learnt about an item, it will be generalised to other associated objects. However, these generalisations can be erroneous so concepts should also change with subsequent experiences that contradict them. A characteristic of Rogers and McClelland's computational theory is that the acquisition and generalisation of feature knowledge is adevelopmental; it ignores cognitive abilities that develop with age. According to Rogers and McClelland (2008) experience, not age, corrects the inaccurate generalisations. As children gain experience, their semantic network matures.

1.5 Relational categories and concepts – later development

An experiential view of conceptual development proposes that direct experience with objects in the environment promotes learning. According to this approach, parental scaffolding, especially labelling, is important. Such experiences can aid children in their discovery of common features of objects (Smith et al., 2002). As children get older, they develop an awareness of the relationships between concepts and categories. For example, they can appreciate taxonomic associations. An emerging sense of taxonomy as a basis for sorting objects is evident in pre-school age children. By 3 years of age, children can sort pictures into the familiar superordinate categories of animals, food and clothing (Waxman & Gelman, 1986). However, providing children with the appropriate category labels had a facilitative effect on sorting; 3-year-olds were more accurate when labels were used than when they were not, implying verbal labelling continues to provide a basis for sorting objects in more abstract categories (Waxman & Gelman, 1986). In a study by Tunnicliffe & Reiss (1999), 5-year-olds could group animals, although mostly according to their anatomy i.e. they chose to sort stimuli according to their perceptual features. Using the sortal of artefacts, Freund, Baker & Sonnenschein, (1990) explored whether 3- and 5-year-olds could sort furniture into more refined categories, such as chairs. It was found that 5-year-olds were more accurate than 3-year-olds. By the age of 5 years, children seem to be aware that the superordinate category of furniture can be broken down into different subgroups, but only in the context of accessing heuristics, such as category labels or perceptual attributes where possible.

Children also tend to orient to basic-level categories i.e. the category that provides a high level of cue validity for identifying the object as a member of the category. In an early study with adults on category verification (Rosch et al., 1976), images at the basic level (e.g. *apple*) were classified faster than images at either the superordinate (e.g. *fruit*) or subordinate (e.g. *Golden Delicious apple*) levels. According to Murphy and Brownell's (1985) differentiation hypothesis, Rosch's finding that the basic level produces better performance is due to the specificity and distinctiveness of objects at this level, compared to the superordinate and subordinate levels respectively. In this way, objects encountered at the basic level are better differentiated with more characterising features. A seminal experiment by Mervis and Crisafi (1982) that used nonsense stimuli found that children can match images from basic, but not subordinate categories, by 4 years of age. Whereas items within each basic level category were similar in shape, with shared features that were not present in other basic categories, there was a large overlap of features between the two subordinate categories. Children were successful at the subordinate matching task by approximately 5-6 years of age. In a related finding, only children aged 4-5 years and older could sort items at the subordinate level (Saxby & Anglin, 1983). The youngest group (aged 3 years) were less able to use the subordinate level as a sortal. There could be several explanations for this. One possibility is that the features of objects in these categories closely overlap with those of concepts in other subordinate categories. Consistent with the claims of the parallel distributed processing (PDP) model (Rogers & McClelland, 2008), 3-year-olds have not yet acquired the features that differentiate subordinate level categories, as a consequence of their short duration of exposure to multiple features. Therefore, the familiarity of concepts with typical features is an important benchmark in early conceptual development, as it then allows the differentiation of subordinate categories (e.g. Pink Lady apple versus Golden Delicious apple).

In parallel to implicit increases in the familiarity of concepts, children also show emerging degrees of explicit insight to conceptual organisation from preschool age. By 4 years of age, children are aware of taxonomic relations since they can justify putting taxonomically related items together (Smiley & Brown, 1979). However, they still maintain a preference for selecting thematic associations. For example, when presented with a probe and two other pictures: one taxonomically and one thematically related to the probe, these children chose the thematically associated image more often. By age 8, children start to group items using explicit reasons based on taxonomy (Tunnicliffe & Reiss, 1999).

Overall, children develop an awareness of taxonomic relations from exposure to verbal, superordinate labels and a bias to construct categories according to their familiarity with perceptual features of objects in their environment. Mastery of discriminating between concepts emerges slowly; by 4-5 years of age children can distinguish concepts that belong to basic level and subordinate categories.

1.6 Category membership and typicality

As reviewed in section 1.5 above, sorting by subordinate category develops from approximately 4-5 years of age (Murphy & Brownell, 1985; Saxby & Anglin, 1983). However, a category verification study provides evidence that adults can struggle with subordinate sorting (Murphy & Brownell, 1985). Adults were slower at categorising at the subordinate (compared to basic) level when items were *typical* of the basic category to which they belonged. An example of a typical object is a goldfish. It is typical because it has many characteristic properties of the basic '*fish*' category (e.g. fins and a tail). Participants were presumed to be faster with category verification at the basic level (e.g. is it a *fish*) because this sortal differentiates concepts more readily than using a subordinate sortal, where the individual concept needs to be directly identified and has fewer distinguishing features (e.g. goldfish versus carp). Paradoxically, 5-6-year-olds did not appear to find subordinate sorting challenging (Saxby & Anglin, 1983). One explanation is that in Murphy and Brownell's

(1985) experiment, participants saw pictures one at a time. In contrast, Saxby and Anglin (1983) gave participants a set of stimuli. Seeing more than one item at once may have had a facilitative effect for shared features that determine the similarity of objects within a sortal, even for subordinate classification. For adults in Murphy & Brownell (1985), the shared similarity across objects at the subordinate level was not available, rather, participants had to directly identify the subordinate category of the image they were looking at, accessing distinctive features. This points to a more complex role of contextual and task demands when evaluating category knowledge in children and adults. This suggests that cognitive processes in later conceptual development deserve more scrutiny. We cannot simply attribute children's performance on semantically-driven tasks to their limited experience and familiarity with features.

The typicality of concepts can influence a range of language processing tasks, including naming (Dell'Acqua, Lotto & Job, 2000) and category verification of objects (Jerger & Damian, 2005). The typicality of an object correlates with naming latencies in Italian adults (Dell'Acqua et al., 2000) with more typical pictures being named more quickly than those that are less typical. Although typicality is an important characteristic of object processing, the relationship to conceptual development is less clear. A recent study reports a correlation between the perceived age of acquisition (AoA) of an object and its rated typicality: early acquired words are generally rated as being more typical (Shröder, Gemballa, Ruppin & Wartenburger, 2012). One AoA study found that early-acquired words are processed faster than late-acquired words by children between the ages of 10 and 18 (Assink, van Well & Knuijt, 2003). AoA effects have also been found in adults when pictures have been used, rather than words (Bonin, Chalard, Méot & Fayol, 2002). One suggestion is that AoA effects occur because the semantic network becomes less flexible as more concepts are acquired (Ellis & Lambon Ralph, 2000). Although there is a correlation between AoA and typicality (Shröder et al., 2012), some low typicality concepts can also be acquired early in development (Holmes & Ellis, 2006). For early acquired items, more category verification errors were made when typicality was low as compared to high, whereas AoA did not have an effect on category verification accuracy or latency. This tentatively suggests that typicality cannot be simply explained away as a function of the age at which a concept is acquired.

1.7 Typicality effects as a window to conceptual development

Various theories attempt to explain the structure of the semantic system. Probabilistic or prototype theory (Medin, Ross & Markman, 2005) proposes that concepts in semantic memory are organised into categories. Properties that characterise many of the members of a group form the basis for organisation. For example, most birds have the ability to fly. We have a conceptualisation of what members of a category are 'generally like' and this is referred to as the prototype for the category. According to prototype theory, novel objects are compared to the prototypical member of a given category and this promotes faster classification if the novel object is more typical of the category, for example due to feature overlap with the prototype. Another approach is Exemplar theory. This suggests that when we see an example of a category we store it in memory (Nosofsky, 1986). For example, we may see a goldfish, which is an example of the *aquatic animal* super-ordinate category. Exemplar theory emphasises the similarity of each new object compared to stored exemplars within a given category. Our decision about whether the new item is part of the category in question is based on how similar it is to our stored exemplars. According to this theory, typical objects are categorised quickly by their similarity to the many exemplars stored in memory. In a mixed model proposed by Medin and colleagues (Medin, Altom & Murphy, 1984), these theoretical approaches are combined so that both prototype and exemplar information guide classification decisions. This model suggests a processing advantage for typical members of categories as many of their features overlap with those of both prototypes and exemplars stored in memory.

To date, few studies with children have investigated typicality effects in object processing tasks. Typicality can influence children's performance, for example, on activities involving recognition of patterns (Posnansky & Neumann, 1976) and the free recall of words (Bjorklund & Thompson,

1983) where children between the ages of 6 and 12 recalled a greater number of typical, as compared to atypical, words. Typicality can also influence older children's ability to match and verify category membership (Duncan & Kellas, 1978; Jerger & Damian, 2005). Children > 8 years were faster to judge whether two images belonged to the same category, if they were typical rather than atypical (Duncan & Kellas, 1978). In a cross-sectional developmental study, Jerger and Damian (2005) found that children as young as four years of age made fewer errors when classifying typical items as compared to atypical ones. Children's reaction times were also faster for the typical objects. Interestingly, age effects were evident with younger children being much faster and much more accurate at classifying typical, as compared to atypical, images as category members.

One issue with this research field is that only highly familiar objects could be suitable for probing children's language processing ability based on their lack of experience of less familiar or typical objects. If there is a strong relationship between familiarity of objects and typicality of their category (Ashcraft, 1978), then choosing items based on their familiarity should bias selection towards more typical members of categories in most research with children i.e. atypical items are not 'known' by young children. Similarly, the limited research on AoA effects in developmental populations could reflect confounds with familiarity or other task-driven factors (Funnell et al., 2006). One study found that children between the ages of 3 and 5 years named images of early acquired words faster and more accurately than those of late acquired words (Anderson, 2008) but this could reflect the increased familiarity of the object. The developmental scope of typicality effects in relation to age has not been fully explored in the identification of object concepts, for example in children's naming and cross-modality matching, or in the differentiation of concepts, such as in category matching, verification and sorting tasks.

One developmental theory of typicality proposed by Jerger and Damian (2005) was that children's early categories <u>only</u> include typical objects. For example, the bird category for 4-5-year-olds may only include magpies (as typical birds), but not atypical penguins. This proposal is consistent with two ideas i) that typical objects are more likely to be frequently encountered and therefore are central to the formation of the category and ii) that categories are graded (or unequal) in terms of the characterising attributes required of an object for category membership. Engaging with atypical objects seems to reflect the advancement of age, although it is not clear whether the affordances of age simply reflect an experiential account of conceptual development, or other maturational factors.

Rogers and McClelland's (2008) connectionist model implements experiential learning using connection weights to drive the most frequent attributes and relations, on the assumption that these pairings correspond to more typical featural attributes. However, by extension, less typical attributes and relations (and therefore atypical objects) are less likely to receive a strong connection weighting due to infrequent exposure or experience, and therefore are acquired more slowly across childhood. In this model, atypical objects can only be accepted as members of the appropriate categories as part of a delayed developmental trajectory of acquisition driven by experience of features. As the features of atypical objects become more familiar, then any advantage for processing more typical objects should attenuate with age.

Others argue that it is not when different types of concepts are acquired, but how we use and acquire concept knowledge that matters (Funnell et al., 2006). In this investigation, children between the ages of 3 years 7 months and 11 years 6 months were asked to name and answer questions about a set of objects. Both naming and knowledge of objects developed with age. However, the developmental change in naming ability did not match the timing of developmental changes in the ability to answer questions about the concepts. It appeared that a qualitative shift occurred in children's semantic organisation at around 79 months. Younger children (below 79 months) exhibited a strong tendency for naming items rather than providing detailed knowledge about objects, whereas older children (aged 79 months or above) exhibited more detailed knowledge compared to their ability to name objects. Funnell and colleagues (2006) argued that their results could not simply be explained in terms of AoA effects (Ellis & Lambon Ralph, 2000). If the semantic system is only shaped by words that are acquired early in development, then the accumulation of names and concept knowledge should slow down at the same rate. Funnell and colleagues (2006) highlighted the developmental context for learning about objects. For example, early acquired words could show an advantage in children's naming because richer visual information is encountered when the words are learned, on the presumption that the item itself is generally present. According to this view, by encoding a richer array of perceptual and semantic features, the input for younger children could influence the cognitive processes that organise and retrieve object knowledge. A key emerging question is, therefore, when and how are children equipped with cognitive processes that support later changes in conceptual development?

1.8 Controlled semantic cognition – a new framework for semantic cognition

Over recent years, a new theoretical approach to conceptual knowledge has emerged, largely based on the neuropsychology and neuroscience profiles of adult patients with brain insult (Jefferies & Lambon Ralph, 2006). This theory is known as the controlled semantic cognition (CSC) framework (Rogers et al., 2015) and incorporates two elements; semantic representation and semantic control. To support semantic representation, the semantic system contains a cross-modal hub and modality-specific spokes. The spokes hold information relating to just one modality and they are connected to the multimodal hub, which is located bilaterally in the anterior temporal cortex (Lambon Ralph, 2014). The hub allows us to identify similarities between items (Rogers et al., 2015) that afford generalisation to other items based on these similarities. Critically, processing semantic features is only one aspect of semantic cognition, the semantic control system ensures the relevant information is activated in each situation or context (Jefferies & Lambon Ralph, 2006). Control requirements are increased in semantic tasks when the options to choose from are similar to one another. Under these conditions, controlled activation is necessary to select the target among competing alternatives (Wilshire & McCarthy, 2002). Control requirements are also greater when the semantic relationship to be identified is weak as compared to strong. When a probe is presented, representations of weak associates are not automatically activated. Therefore, controlled retrieval is required to regulate activation of semantic knowledge (Hoffman et al., 2018). Controlled retrieval is a top-down mechanism that boosts activation of weakly related objects when bottom-up processing cannot identify the target.

Two contrasting neuropsychological profiles have contributed significantly to the development of the CSC framework (Rogers et al., 2015). These are those of individuals with a diagnosis of a progressive neurodegenerative disorder referred to as Semantic Dementia (SD) and individuals with brain insult, largely through stroke, referred to as Semantic Aphasia (SA; Corbett, Jefferies, Ehsan & Lambon Ralph, 2009). In both conditions, people have difficulty with comprehension tasks that assess semantic knowledge using words, pictures and stories. In a seminal study that compared these two patient groups directly, Jefferies and Lambon Ralph (2006) administered a set of semantic tasks that draw on access to conceptual representation through object identification (e.g. spoken naming and word-picture matching) and object discrimination and association (e.g. picture and word association tasks). Notably, the task demands were also the focus of this study. For example, spoken picture naming demands that participants produce the name of the object they see. This requires specific knowledge of object identity because the correct name must be selected from a pool of competitors. Individuals that do not have specific knowledge of identity will not be able to discriminate between semantically-related items. For comparison, picture matching tasks, such as the camel and cactus test (CCT), were also used. In the CCT, participants choose which of four items is related to a target. For example, they may have to decide whether *cactus, tree, sunflower* or *rose* is associated with *camel.* To be successful on these tasks, participants must have broad associative knowledge that relates thematically to probe items. For patients with SD, semantic tasks were uniformly impaired suggesting a consistent loss of concept knowledge. Item correlations were also established across tasks using different modalities to assess awareness of associative semantic relations, suggesting consistency for individual concepts. It was concluded that their conceptual difficulties could be attributed to the degradation of amodal semantic representations. For patients with SA, semantic tasks were not uniformly impaired across tasks, or to the same degree as patients with SD, and the pattern of correlations differed across tasks. Although picture and word versions of associative matching were correlated, their scores for associative picture matching and word-picture matching were not correlated. Jefferies and Lambon Ralph concluded that these SA patients were not simply showing evidence of conceptual impairment. The SD patients had degraded amodal conceptual representations, since changing the semantic task and modality did not alter performance so long as the same item was being tested. By contrast, the SA patients showed a pattern of difficulties that aligned with the variable role of cognitive control in the semantic tasks. Specifically, the control requirements of the associative picture matching task were greater than the demands of word-picture matching tasks that captured object identity. The lack of specific relations on associative matching tasks places greater cognitive demands for selecting the appropriate match, whilst rejecting alternative semantic associates. From this perspective, it was argued that SA patients experience difficulties on semantic tasks because they cannot direct and use semantic information in a task-appropriate fashion. SA patients struggle when selection demands are high and when there is a requirement for controlled retrieval. Healthy older adults (61-91-year-olds) have similar problems resolving selection demands (Hoffman, 2018), despite having more breadth of knowledge than younger adults (18-30-year-olds) and fewer difficulties with controlled retrieval.

Supporting evidence for impaired semantic control in SA patients comes from a study of object-function relations in this group (Corbett et al., 2009). In this study, an object-use battery using verbal and non-verbal tests was administered. Participants had to match words with pictures of those items i.e. capturing object identity. They also had to match items according to specific attributes, such as the object's function and the recipient of the action. The SA patients scored higher on word-picture matching than object-function and object-recipient matching tasks. The authors proposed that this was because the control requirements of the associative matching tasks were greater than those of the word-picture matching task (Corbett et al., 2009). The underlying argument of the CSC framework is that, although experience is important for learning about concepts and their features and relations to one another, it is not the only element involved in semantic cognition. Other factors, such as the ability to select from a range of associated representations, are also crucial. The availability of related concepts can be a hinderance because it can result in more options to choose from and therefore increase the demands of selection. For example, when the extrinsic requirement of the task is to identify a single object, this process may be slowed down if many items with similar features or thematically related concepts are known and brought to mind.

1.9 Typicality and task demands in the controlled semantic cognition framework

Although early theoretical approaches emphasised the frequency, as well as typicality, of objects as having a beneficial effect on processing of typical objects, the CSC framework does not. According to the CSC framework, typical objects should activate the representations of multiple related objects within a common category, due to the overlap of features between the probe and other members of the same category. The automatic activation of shared features increases selection demands that challenge the ability to identify the object that is most relevant to the task at hand. For highly typical objects, there is greater competition between activated semantic neighbours that poses a risk to the correct selection of the relevant object concept in identification tasks. For tasks that rely on discrimination of objects by common features, such as sorting tasks, the automatic activation of related concepts does not hinder performance to the same degree. When identifying the category that the probe belongs to, all activated representations belong to the same category; there is no relative disadvantage for typical objects due to automatic activation of shared features.

One recent study by Rogers and colleagues (2015) investigated the effects of typicality, familiarity and specificity of objects in relation to performance on measures of object identification and object association in SA and SD patients. This involved four semantic tasks: naming, word-picture matching, sorting and category fluency. Two subsets of pictures were used: one was a 'typicality' subset made up of 16 triplets, which were matched for rated familiarity. One of the pictures in each triplet was highly typical, one was moderately typical and one was atypical. The second subset of pictures was a 'specificity' subset. This subset contained 22 pairs of items. Both pictures in a pair came from the same intermediate/basic-level category, but one had a higher familiarity rating than the other. Rather than simply estimate the degree of impairment across tasks and patient profile, the study aimed to see if the profile of performance across groups met the expectations of the controlled semantic cognition framework. Findings from this study will be discussed in sections 1.9.1-1.9.3.

1.9.1 Typicality effects in naming tasks. Naming tasks require specific identification of the object through a common label, typically named at the subordinate level for familiar objects. Naming tasks are susceptible to a range of psycholinguistic variables that can affect word production tasks, such as familiarity, imageability, concreteness, AoA (Bonin et al., 2002), frequency (Barry, Morrison & Ellis, 1997) and operativity of the item. Operativity refers to whether the item can be manipulated or not. Aphasic patients have been found to be more accurate at naming operative and high frequency pictures (Gardner, 1973). Imageability also has a significant effect on naming in patients with aphasia (Nickels, 1995, as cited in Nickels & Howard, 1995). Finally, children are faster and more accurate at naming pictures of words that are acquired early as compared to late (Anderson, 2008).

In their naming task, Rogers and colleagues (2015) predicted that the SD patients would benefit from naming more typical compared to less typical objects. Although these patients have degraded knowledge of the properties that differentiate objects, typical concepts are more robust because they do not possess many differentiating properties compared to close neighbours. In contrast, atypical objects have more individuating features that are crucial for identifying the object and are therefore susceptible to weaker retrieval. By contrast, the SA patients (and controls) would not benefit from naming more typical objects due to the degree of automatic competition generated amongst category members similar to these objects. Without sufficient executive control to overcome these selection demands, naming of highly typical items would be vulnerable to weaker selection of the correct target, therefore removing any performance advantage for typicality. The data reinforced this pattern; there was a benefit for naming typical items in the SD group, but no typicality effect in either the SA group or controls. In fact, the controls showed a reverse typicality effect, arguably because selection demands increased for these more typical pictures and were detrimental when the precise identity of the probe was required. Further supporting evidence for the presence of typicality effects in patients with SD was provided by Patterson, Kopelman, Woollams, Brownsett, Geranmayeh & Wise (2015). However, the absence of typicality effects in the SA group is not consistently observed. Rossiter and Best (2013) investigated the effects of typicality on naming in patients with acquired aphasia. High and low typical items were matched for familiarity, imageability, concreteness, AoA, operativity, length and frequency. Rossiter and Best found that patients with aphasia were better at naming more typical objects than atypical objects. This variation was perhaps due to the selection criteria across studies. In the study by Rogers and colleagues (2015), patients were only selected if they experienced problems with word completion and the picture version of a difficult four alternative forced choice semantic association test i.e. they favoured selecting individuals with semantic deficits, not broader attributes of aphasia. Rogers and colleagues argued that the typicality advantage in naming did not occur in their SA patients because this group could not resolve the selection demands generated by competition between typical items and their semantic neighbours.

1.9.2 Typicality effects in item matching tasks. Word-picture (item) matching tasks require participants to match a spoken word with a picture of the exact same object, but where the target is presented alongside distractors. Rogers and colleagues (2015) investigated typicality effects in word-picture matching tasks by manipulating the semantic relatedness (distance) between the distractors and the target. By using subordinate taxonomic relations between the picture target (e.g. *cheetah*) and distractors (e.g. *leopard, tiger*), the distractors occupy the same subordinate category and evoke high levels of featural similarity (has *whiskers*) and relational similarity (can *hunt*) to the target item. In this case, the competition from distractors is high because they share proximal semantic relations with the target item. Therefore, competition for selection is greater for typical items due to the degree of overlapping and shared relations amongst other category members.

For the word-picture matching task, Rogers and colleagues found that the SD and SA patients were more accurate at matching <u>more typical</u> (as compared to less typical) items when distractors were distal. However, semantic patients were also more accurate at matching <u>less typical</u> (as compared to more typical) items when the distractors were proximal. They argued that both patient groups could not overcome the task demands for selecting the correct picture, in the presence of distractors. If the distractors shared a proximal relation to the target, then both groups of patients were more impaired on the typical items, due to the increased level of competition between the number of shared features and shared relations invoked by the typical target compared to distractors.

1.9.3 Typicality effects in sorting tasks. Sorting tasks require objects to be differentiated from one another and grouped according to a sortal that reflects category membership. The difficulty of the task depends on the shared semantic relationship between the target object and the sortal. Typical objects share lots of common semantic and featural relations with category members, however the degree of automatic competition is higher when the sortal is more specific to the category. For example, if the sortal is '*living things*', then a bird (whether *robin or penguin*) is not more relevant or typical. However, if the sortal is more specific (e.g. *birds)*, then the differentiation between robin and penguin can be relevant. This is because the shared features and relations of a typical bird (e.g. *robin*) are higher than those of the less typical one (e.g. *penguin*). According to the CSC framework, Rogers and colleagues (2015) suggest that sorting atypical (rather than typical) objects into specific categories requires greater semantic control. This is because the varied features and semantic attributes of the atypical objects also activate members of alternative categories, increasing the selection demands of selecting the correct category. For this reason, the level of the sortal term is important. For example, under a *specific* sortal term where subordinate category options are closely related, then an activated representation from an alternate cluster could result in selecting the incorrect category label. In this way, participants can succumb to selection demands and fail to select the cluster that the probe belongs to. Semantic control is required to dampen the semantic relations engaged by the activated item brought to mind, to enable a correct selection. Since typical items, as compared to atypical ones, may benefit from being more strongly associated with their cluster, then the automatic activation of this category relation is sufficient to enable a correct selection. By contrast, an atypical probe increases the need for controlled cognition due to the increasing activation of semantic relations from more distal clusters.

Rogers and colleagues (2015) investigated typicality effects in two sorting tasks. SA and SD patients were required to sort typical and atypical pictures into three semantic categories. The sortal instructions varied in terms of specificity; they were either very general categories (plant, animal, manmade object) or more specific categories. For example, in one of the specific sorting tasks participants had to decide whether animals were found on land, in the air or in water.

For the specific sorting tasks, Rogers and colleagues found both SA patients and SD patients were more accurate at categorising typical objects as compared to atypical objects. However, typicality did not have an effect on performance in the general sorting task. These findings are consistent with the CSC framework (Rogers et al., 2015). Patients may have been less accurate with atypical items in the specific sort because they brought to mind members of alternative categories. Typicality was not expected to have as much of an effect on general sorting because, in this task, the relationship between the categories was more distal. Therefore, even if atypical images activated representations from different categories that were closely related to the target category, these were unlikely to be options that the participants could choose from so selection demands were not increased.

1.10 Exploring the controlled semantic cognition framework in children

1.10.1 Evidence for contextual demands in semantically-driven tasks.

The experiential view is an important perspective that underpins most of the theorising about children's conceptual development. However, some results challenge the view that acquired experience (knowledge) of concepts is a sole determinant of semantic cognition. If an individual has detailed knowledge of concepts, it is highly likely they will be successful on semantically-driven tasks involving these concepts. Yet, the influence of task demands and the interaction with complex cognition can also be important. For example, presenting stimuli as a set, rather than individually can have a facilitative effect on subordinate sorting (Saxby & Anglin, 1983). The CSC framework (Rogers et al., 2015) offers a re-interpretation of behaviour on semantically-driven tasks. For example, that sorting of subordinate categories induces automatic activation of featural detail and relations within categories. The control demands for achieving a correct sortal are reduced when items are present in the array and, therefore, the task demands are not relying on controlled retrieval processes to the same extent. For younger children, category labels can be helpful for general sorting (Waxman & Gelman, 1986) due to the activation of similar and related category members being brought to mind, increasing the likelihood of more typical category members being retrieved as an aid to sorting. It is likely that the CSC framework (Rogers et al.,

2015) can provide a richer interpretation of the development of children's performance on semantically-driven tasks.

1.10.2 Conceptual flexibility in children

One line of research with children has probed 'categorical flexibility'; the ability to switch between different representations of the same item (Blaye & Jacques, 2009). We can think of objects in terms of their taxonomic relations but also their thematic associations. The study by Blaye and Jacques (2009) required 3-5-year-olds to match probes to a related target based on categorical and thematic relations to the target. Children had to choose a match for a probe from three options (a thematic associate, a taxonomic associate and an unrelated item). After their first selection, they were required to select a further match from the two remaining pictures, with the third picture present. In doing so, the task required flexibility in switching from one semantic relation to another for a correct match, tapping into general executive processes. Five-year-olds performed better than 4-year-olds when calculating the proportion of correct double selections overall, implying that children's aptitude for selecting semantic relations (not just response selection) improves with age. Further evidence from Deák (2000b) probed children's ability to switch from thinking about different semantic relations of a common object, e.g. what something 'looks like' to what it is 'made of'. In this task, children were told a fact about a standard object. For example, it "looks like an X" (where X was a novel word). The child then had to choose the matching object from four options. One was a foil and the other three targets shared one feature with the standard. Children aged 3 years were able to select the object that also 'looked like an X' and performed above chance. However, when the semantic relation was switched, for example to 'what this one is made of' in further blocks, then only older 4- and 6-year-olds performed above chance. Younger children were only able to choose options that were appropriate for facts that they'd heard on previous trials, implying a lack of conceptual flexibility. Together, these studies have been argued in favour of a general developmental improvement in executive cognition that can influence the selection of semantic relations around four-six years of age. To date, developmental research has focussed on younger children's ability to explicitly acknowledge and select different semantic relations of the same object. Despite evidence of 'conceptual flexibility' in 4-5-year-old children that is linked to general executive skills, it is not clear whether children, like adults, experience competitive activation of related concepts in the form of implicit selection demands on semantically-driven tasks.

To date, two studies have explored whether children are able to resolve selection demands (Snyder & Munakata, 2010; Snyder & Munakata, 2013). Providing subcategory labels in a verbal fluency task increased the number of words generated by 4-5-year-olds (Snyder & Munakata, 2010). By introducing subcategory labels, any selection demands are reduced since the labels guide children to select from categories with fewer members. Without subcategory labels, children were required to select from many options, impeding their performance. In a further experiment with 6-year-olds (Snyder & Munakata, 2013), children had to name items repeatedly in a blocked cyclic naming task. Response times were faster for mixed blocks (where all pictures were from different categories) than for homogeneous blocks (where images were from the same category). Competition is thought to be high in homogeneous blocks because activation spreads to semantic neighbours (Schnur, Schwartz, Brecher & Hodgson, 2006). This means that executive processes are required in order to select the correct name. Overall, 4-6-year-olds have difficulties when selection demands are high, suggesting weaker recruitment of controlled cognition. Older children may have more mature cognitive control systems that could enable them to overcome these demands. However, the involvement of general executive processes, or processes specific to semantic retrieval remains unresolved.

1.11 Aims and objectives

The study aimed to identify if, and when, controlled semantic cognition emerges in childhood to support goal-appropriate retrieval of concepts. It is likely that controlled semantic cognition in children depends on two aspects a) the availability of a rich network of related concepts and b) the engagement of cognitive control processes to guide semantically-driven behaviour. Following evidence of qualitative changes in the regulation of semantic knowledge on different tasks at 6 years 6 months (Funnell et al., 2006), it was hypothesised that managing selection demands in a rich and clustered semantic network engages controlled cognition at around 6-7 years of age, with more tentative engagement of controlled cognition in younger children below 79 months, as noted in the findings of Snyder and Munakata (2010).

One way to identify controlled cognition is to manipulate task requirements for semantic decisions about a common set of objects. Following the design of Rogers and colleagues (2015), performance on four semantically-driven tasks (picture naming, item (word-picture) matching, sorting and relational matching by category) was compared using the same object array. This approach implies that any variation across tasks relates to the task requirements, rather than variation in the availability of specific objects as 'known' or 'not known' to the child.

A second way to identify controlled cognition was to manipulate the typicality of objects in this set. Graded typicality effects usually refer to a positive endorsement of more typical items within a given task. Typicality effects are usually interpreted as due to the benefit of more prototypical features, or more familiarity of the objects in tasks. For children, graded typicality effects have been observed in pattern recognition (Posnansky & Neumann, 1976), free recall of words (Bjorklund & Thompson, 1983), category matching (Duncan & Kellas, 1978) and category verification (Jerger & Damian, 2005). However, Rogers and colleagues (2015) characterised the interaction between task requirements and typicality effects, so that more typical objects arouse more activation of shared semantic relations and features with their same-category peers, that can be detrimental to performance, if familiarity is controlled.

A final way to illustrate the influence of controlled cognition was to manipulate the semantic relatedness of distractor items that influence the response selection on matching tasks. Three matching tasks systematically varied the type of semantic relation of distractors to target i.e. cross-modality matching by identity, relational matching by semantic category and relational matching by object-use. By manipulating the semantic relation between the target item and the distractors, through distal or proximal semantic relations, the requirements for controlled cognition were manipulated across tasks.

1.12. The present study – overview

To investigate controlled semantic cognition in school-age children, five semantically-driven tasks were administered to participating children aged between 5-10 years. Three tasks (picture naming, picture sorting and item matching) were derived from a semantic battery presented to adults with and without neuropsychological impairment (Rogers et al., 2015). Here, the tasks were adapted for use with children in two ways i) by providing a simple narrative in which children had to help a forgetful cartoon character, Peter the Panda and ii) by altering the number of distractors in the array for matching tasks. The predictions for performance on each task are discussed below.

1.13.1 Predictions for the picture naming task

Naming requires specific knowledge and retrieval of the identity target item and, therefore, requires mastery of selecting from related semantic neighbours to support the mapping of the semantic concept to its spoken form, in addition to word production processes. Since the target items were selected as familiar and common objects for children, the experiential view predicted little improvement in picture naming with increasing age, since the objects were likely to be known and familiar to all age groups **i.e. small to negligible effects of age**. A more specific claim is that children learn facts about typical objects before they gain information about atypical items (Jerger

& Damian, 2005) implying more familiarity with more typical objects. Since familiarity was controlled across typical and less typical items in the current study, variation in object knowledge by experience would be reduced. If there is a developmental lag in the acquisition of atypical items by experience, then children would be predicted to produce typical words more quickly **i.e. a main effect of typicality on picture naming latencies.**

The CSC framework (Rogers et al., 2015) suggests that typical pictures bring to mind semantic neighbours that serve to increase competition amongst shared properties of items for retrieval. Adult controls were found to be more accurate at naming atypical items, when controlling for familiarity. For older children, any influence of semantic competition on typical items was more likely to emerge after 78 months, consistent with the recruitment of controlled semantic cognition around this age. Therefore, older children > 78 months were anticipated to become more prone to the competitive selection demands of typical concepts **i.e. an interaction of age*typicality with more accurate and faster naming of less typical items in children > 78 months.**

1.13.2 Predictions for the item (word-picture) matching task

Word-picture matching tasks for individual items emphasise the identification of a single object across different modalities that both tap into a single multi-modal construct for that object. Successful matching depends on children's word knowledge and the ability to retrieve relevant distinguishing features. Word-picture matching is a standard measure of the child's breadth of knowledge about single constructs (see Laws, Briscoe, Ang, Brown, Hermena & Kapikian, 2014 for related discussion in atypical development). Since the probe items were selected as familiar and common objects for children, the experiential view predicted little improvement in item matching with increasing age, since the objects were likely to be known and familiar to all age groups **i.e. small to negligible effects of age**.

Item matching and typicality of objects. Item matching requires access to a set of distinguishing features for identification. More typical objects share semantic properties with other related category members. However, an experiential view predicts that children's item matching would be insensitive to the distinctiveness or typicality of the target item, once the item is familiar and known.

The CSC framework incorporates children's ability to draw on controlled cognition to direct semantic retrieval. In the present design, distractors were manipulated to exploit any semantic relations between the target item and distractors with shared semantic properties. Proximal distractors shared the same category as the target, implying that featural relations would be more likely to be shared compared to distractors in the distal condition, since the sub-categories of these distractors did not overlap with that of the target. Overall, the proximal distractors should induce the need for more controlled cognition, when semantic relations are overlapping with the target picture. Furthermore, if objects are more typical of their category, then they should require more controlled cognition when distractors are proximal, rather than distal. By comparison, if objects are atypical of their category, then the requirements of controlled cognition should increase in the distal condition, where there is overlap with the semantic properties of items from other categories. This pattern should be reflected in the accuracy of item-matching scores in a **typicality*condition interaction.**

As children aged < 79 months were presumed to be less able to recruit cognitive control, then a further interaction by age was likely when matching typical objects in the proximal condition, where the distractors share more semantic properties with the target. Compared to older age groups, the youngest age group was less likely to be able to combat the effects of distractors on typical items **i.e. an age*typicality*condition interaction**.

1.13.3. Predictions for the sorting by category task

Sorting pictures of objects requires either general knowledge of the category to which the objects belong or, depending on the subcategory, more specific knowledge of semantic properties that determine a class of similar objects and allow differentiation of similar semantic neighbours. When sorting on general semantic class in the present study, it was likely that all children would be able to sort typical items across global taxonomies such as animals, plants and vehicles. To identify the categories of typical objects requires knowledge of natural kinds (animals, plants) and common artefactual kinds (vehicles) that develops by preschool age (Gelman, 1988). Younger children may struggle to classify atypical items because their categories are limited (Nelson, 1974). As children get older their categories expand to include atypical objects. The use of subcategories for sorting (such as land, air and water) involves similar objects that share more specific semantic properties and features. Older children will have accumulated more experience of specific and general categories and their semantic properties.

Sorting and typicality of objects. More typical objects are more predictable members of specific categories due to the preponderance of shared semantic properties that characterise the category. By controlling for familiarity across the typical and atypical items, the availability of semantic properties would be similar across item sets. However, if there is a developmental lag for assigning atypical items to either global or specific categories, then the younger children would be less likely to sort the atypical items correctly, compared to older children. According to an experiential view, **an interaction of age and typicality** was anticipated on both **general and specific sorts**.

According to the CSC framework, sorting under general superordinate categories involves more distant taxonomic relations that could recruit more controlled cognition to access the correct category. If younger children have reduced access to controlled cognition, then they are more likely to struggle with the use of a generic sortal term that relies on distal relations, compared to more guided sorting using specific sortal terms. For this reason, **an interaction of age*specificity of sortal term** was also predicted from the CSC framework.

Typical items can bring to mind other items from within the same specific category, but this activity would be likely to enhance the selection of the correct category. By contrast, atypical objects can activate representations of items from alternative categories that compete with the correct category choice. Under a general sortal of a superordinate category, it would be unlikely to class an atypical target under either of the incorrect categories. Under a specific sortal of a subordinate category, it would be more likely to class an atypical target under one of the incorrect categories. Since atypical items bring to mind members of other categories, increasing the competition across alternative categories in specific sorting, then **a specificity*typicality interaction** was predicted. Since the youngest children would be less able to recruit controlled cognition to resolve the competition between categories for the atypical items, then an advantage for sorting typical items under a specific sortal would be more pronounced in this age group i.e. an **interaction of typicality by age** for the specific sort task.

1.13.4 Predictions for the relational matching by category task

Category picture-picture matching tasks draw directly on the semantic properties of the objects and their relations to each other within categories. Successful category matching depends on children's knowledge of category members, their semantic relations and an awareness of the 'kind' of concepts related to each other within categories and kinds. For example, the feline features of a leopard may be accessed through the semantic similarity to the features of a lion, without direct knowledge of, or explicitly labelling 'feline' as a category. Poor category assignment by younger children could predict slight age variation in matching accuracy **i.e. small age effect** on category matching.

Category matching and typicality of objects. More typical objects can be defined through their semantic relations to other members of a common sub-category. Accessing semantic properties and relational information is essential to make a competent match to members of the same category (e.g. similarities between lions and tigers). Despite the familiarity of the stimuli, it is possible that a lack of enriched category knowledge could affect perceived relations between atypical members of a given category, when matching by semantic features. For example, matching [*jellyfish*] to [*shark*] may lack support from category knowledge that both objects are kinds of sea-based animal even though there is minimal overlap of their distinguishing features. If younger children are disadvantaged by a lack of category knowledge for the atypical items, then this predicts **an age*typicality interaction effect on category matching accuracy.**

In the present study, task demands were manipulated by varying the type of category relation between the probe-target pair and distractors. According to the CSC framework (Rogers et al., 2015), controlled cognition is recruited to direct retrieval and select from competing semantic representations, whether internally or externally generated, to guide successful matching to the target. If the probe object is typical, then proximal distractors elicit more direct competition with the semantic properties of the target object. If the probe object is atypical of its category, then semantic properties are more likely to be shared across alternative categories, creating opportunity for competition in the distal condition. This predicts that matching accuracy depends on the type of distractors and the typicality of the probes, eliciting a **typicality*condition interaction.**

For older children, the recruitment of cognitive control should exploit cues to category membership to guide accurate matching by category. However, younger children are presumed to be less able to exploit controlled cognition, and therefore predicted to be more error prone than older children, particularly when matching probes that require more controlled processing to assign category relations to the target, under conditions of distraction. If younger children (below 79 months) are less able to recruit controlled cognition, then matching errors are predicted to be more evident when matching typical items in the proximal condition and when matching atypical objects in the presence of distally-related distractors, therefore predicting **an overall age*typicality*condition interaction.**

1.13.5 Predictions for the relational matching by object-use task

To be successful on object-use matching tasks one must possess associative knowledge relating the specific attribute of the function of the item (what it is used for) to items that are in receipt of the action. An experiential approach to development would suggest that this implicit knowledge of object functions emerges through practical experience with objects and from linguistic experience that develops sensitivity to the object of a noun phrase (Gerken, Wilson & Lewis, 2005; Wright Cassidy & Kelly, 2001). In the current task, children were asked to match the probe object to another object that would be an appropriate recipient of the functional action of the probe. Critically, distractors were designed to manipulate the proximity of the relationship with the target recipient in two conditions of the task. In the distal condition, the distractors share few semantic properties with the target object i.e. when *pencil* was the target, the distal distractors were *guitar, keyboard, goal* and *wardrobe.* In the proximal condition, more semantic properties were shared between the distractors and the targets e.g. proximal distractors were *rubber, pen, ruler* and *crayon* for the target *pencil.* A prediction based on the experiential view is that children of all ages would achieve high levels of accuracy on this task, regardless of distractor type. Children would make few errors because by five years of age they have developed the appropriate associative knowledge of the functions of the probes and what is necessary for the activities. By contrast, the CSC framework emphasises the role of task demands invoked by the semantic properties and relations between distractors and target objects. It requires greater semantic control to select the target when the distractors are proximal, rather than distally related to the target (Rogers et al., 2015). The proximal distractors elicit competition with the target object because they share many semantic properties with it. In the distal condition, the shared semantic properties are minimal and so distractors provide little competition. Fewer errors were expected in the distal condition as compared to the proximal condition **i.e. a main effect of condition.**

If the availability of semantic control processes that support matching is a correlate of children's emerging general cognitive control, then younger children would be less able to recruit controlled cognition to resolve selection demands for matching the correct target under a semantic load of the distractors. Older children were expected to make fewer errors because they have developed the ability to inhibit distracting information and therefore overcome selection demands of the task. These effects were predicted to be most pronounced for the youngest age group when distractors were proximal **i.e. an age*condition interaction.**

2 Method

2.1 Participants

107 children aged between five and ten years (*Mean age* = 7 years 6 months, *SD* = 1 year 10 months, Age range = 5 years 10 months) participated in this study. Ethical approval was given from the University of Bristol's Faculty of Science Human Research Ethics Committee. Informed consent was obtained from participating families.

Children who were monolingual, native English speakers were recruited from primary schools in a small provincial town in the UK. The majority of children (87.85%) were White British, 4.67% were British-Mixed and 1.87% were White Swedish. British-Indian, Australian-Japanese, New Zealander and White Irish each made up 0.93% of the sample. Ethnicity information was unavailable for 1.87% of the children as their parents did not wish to disclose it. Participants were sampled from schools with a relatively low number of pupils supported by the pupil premium and good levels of scholastic attainment according to school inspection reports (Ofsted, 2004; Ofsted, 2014).

Children were sampled from three age groups. There were 60-78-month-olds ($N = 38$, $Mean\ age_{months} = 67.27$, $SD = 6$), 79-95-month-olds ($N = 35$, $Mean\ age_{months} = 85.61$, $SD = 6.05$) and 96-130-month-olds (N=34, $Mean\ age_{months} = 117.74$, $SD = 9.48$).

Table 1: Standardised scores for language comprehension and non-verbal ability

Descriptor	Age (months)		
	60-78	79-95	96-130
Raven's Coloured Progressive Matrices	101.32 (13.59)	102.71 (16.69)	104.26 (15.77)
Semantic Decisions	9.32 (2.81)	10.14 (3.14)	10.94 (2.26)
Sentence Comprehension	9.45 (3.27)	10.97 (2.51)	10.24 (2.28)

Note. Standard deviations are in parentheses.

2.2 Measures

2.2.1 Basic language comprehension. Two subtests of the Assessment of Comprehension and Expression 6-11 (Adams, Cooke, Crutchley, Hesketh & Reeves, 2001) were used: the Sentence Comprehension and Semantic Decisions subtests. In the Sentence Comprehension subtest, the experimenter read sentences aloud to the child (e.g. The helicopter flew above the clouds) and they had to point to the picture that matched the sentence. The Semantic Decisions subtest involved choosing a synonym for the probe from four options. Each pupil was given a raw score for each measure and a standard score was generated from test norms.

2.2.2 Non-verbal ability. Sets A, A_B and B of Raven's Coloured Progressive Matrices (Raven, 2008) were administered to all children. Children are required to complete a pattern by pointing to the piece that fills the gap or they are given a sequence of patterns and indicate which piece goes in the final slot. The raw score was calculated for each measure with a standard score calculated from test norms.

2.2.3 Semantic cognition battery of tasks. All participants completed five measures of semantic cognition. Three semantic measures (picture sorting, word-picture matching and naming) were adapted for use with children from the Levels of Familiarity, Typicality and Specificity (LOFTS) semantic battery reported by Rogers and colleagues (2015). The forty-eight items of the typicality subset from the LOFTS battery (see Appendix A) were used for four semantic measures: naming, item (word-picture) matching, sorting and category matching. A novel set of 48 colour photographs was generated for use in all semantic tasks, apart from the object-use task. Thirty-six of the photographs were downloaded from online websites, one from the Hatfield Image Test (Adlington, Laws & Gale, 2009) and the remaining 11 were generated directly by the author. Pictures were edited using GIMP-2.8 (Kimball & Mattis, 2012). The photos were adjusted so that they were all similar sizes (approximately 6 x 7cm). The maximum dimension of each picture was 283 pixels and the aspect ratios were kept the same. All pictures included a white background for use in the item matching, sorting and category matching tasks. In the naming task, background context was retained for some items (pheasant, goldfish, aeroplane, canoe, hovercraft, swan, shark, crocodile, submarine, woodpecker, parachute, trout, oil tanker, aircraft carrier, rowing boat, seahorse and helicopter). Images were displayed on a laptop that was positioned directly in front of the child.

Since the set of photographic items was novel, ratings of typicality and familiarity were derived from adults in a pilot study of undergraduate students. Adult typicality ratings were used because adult and child typicality judgements are correlated (Bjorklund, Thompson & Ornstein, 1983). Additional ratings of frequency were obtained using the SUBTLEX-UK database (Van Heuven, Mandera, Keuleers & Brysbaert, 2014). The ratings we used were derived from subtitles for children's TV programmes.

2.3 Ratings of photographic stimuli: All photographic stimuli based on items selected from the LOFTS battery were administered to adult participants, prior to use in the semantic cognition battery. No ratings were acquired for the items used in the object-use matching task.

2.3.1 Participants – ratings study: Fifty undergraduate students who were proficient speakers of English as a first language (*Mean* age = 20 years, 3 months, *SD* = 5 years, *Range* = 35 years) were recruited to a pilot study in exchange for course credits. Ethical approval was gained from the University of Bristol's Faculty of Science Human Research Ethics Committee. Informed consent was obtained from participants.

2.3.2 Materials – ratings study: Undergraduate students were required to rate all 48 photos that had been generated for use in four of the semantic tasks (naming, item matching, sorting and category matching). A list of the selected items is provided in Table 2.

2.3.3 Procedure – ratings study: The task had three parts: naming photographic pictures, then rating pictures for their typicality and familiarity. All photographic stimuli (n=48) were viewed in a single fixed random order within each part of the task. The instructions to participants were based on those given by Rogers and colleagues (2015) in their ratings of typicality and familiarity research. First, participants were asked to name the items as follows: "*We would like to know if our photos are recognisable. Please name the following items".* Second, participants were asked to rate the items for typicality as follows: *"You will see pictures with their category names above them. The pictures will appear one by one. Please rate the items for typicality, where 1 is the most atypical and 7 is the most typical. If you think the object does not belong to the category presented please give it a score of 1. A typical item is something that is a good example of the category it belongs to. For example, a crow can be seen as a typical bird. Please ensure that you use the full range of responses".* In the final part, participants were required to rate the items for familiarity according to the following instructions: *"Please could you now rate the same items for familiarity (where 1 is not familiar at all and 7 is the most familiar). Objects are considered to be familiar if we see them frequently. Please ensure that you use the full range of responses".*

2.3.4 ***Results – ratings study:*** The three parts of the study were analysed separately. For the naming study, a list was compiled of all names provided by at least two adult participants (see Appendix B). Names that had only been given by one participant were removed. In the Naming task, if a child produced a name for an object that appeared as an item on the adult list, it was marked as correct. This is consistent with the procedure adopted by Rogers and colleagues (2015).

A mean typicality rating was calculated for each of the items and they were reclassified accordingly. They were divided into three groups but the new groups were based on our ratings rather than the ones collected by Rogers and colleagues (2015). The new arrangement of objects is shown in Table 2. A one-way independent Analysis of variance (ANOVA) showed that the groups differed significantly in terms of typicality ratings, $F(2, 45) = 64.77$, $p < 0.001$. Contrasts revealed that the ratings for the high typicality group were significantly larger than those for the medium typicality group, $t(45) = 3.45$, $p < 0.001$ (one-tailed) and that ratings in the low typicality category were significantly smaller than those in the medium typicality category, $t(45) = -7.67$, $p < 0.001$ (one-tailed).

Another one-way independent ANOVA revealed that the frequency ratings from the SUBTLEX-UK database (Van Heuven et al., 2014) did not differ significantly across groups, $F(2, 45) = 0.26$, $p > 0.05$.

A final one-way independent ANOVA showed that familiarity ratings from our undergraduate participants did vary according to typicality subset, $F(2, 45) = 4.30$, $p < 0.05$. Contrasts revealed that familiarity ratings did not differ for medium as compared to high typicality pictures, $t(45) = -1.11$, $p > 0.05$ (one-tailed). Familiarity ratings were also similar for low as compared to medium typicality pictures, $t(45) = -1.80$, $p > 0.05$ (one-tailed). The significant difference was between familiarity ratings for low and high typicality images, $t(45) = -2.91$, $p < 0.05$ (one-tailed). High typicality items were rated as more familiar than low ones. However, subjective familiarity ratings have previously been criticised (Brysbaert & Cortese, 2011). Their influence depends on the strength of the objective frequency count used. Therefore, typicality was considered a stable metric, and not a proxy for the familiarity of the items used.

Table 2: Typicality ratings for the photographic stimuli

Typicality					
Low		Medium		High	
Canoe[1]	4.62	Kangaroo[3]	5.90	Chimpanzee[1]	6.18
Shark[8]	4.44	Deer[4]	5.90	Frog[8]	6.08
Ray[8]	4.56	Crab[1]	5.80	Goldfish[1]	6.18
Hedgehog[1]	5.28	Hippopotamus[1]	5.50	Aeroplane[1]	6.08
Penguin[1]	4.78	Kingfisher[5]	5.98	Magpie[8]	6.16
Mouse[1]	5.28	Elephant[1]	5.80	Robin[1]	6.28
Tortoise[8]	5.16	Squirrel[1]	5.36	Trout[1]	6.16
Parachute[1]	3.76	Donkey[1]	5.86	Swan[8]	6.06
Bat[3]	4.54	Duck[4]	5.50	Tiger[1]	6.00
Hovercraft[7]	4.68	Badger[1]	5.48	Cheetah[1]	6.04
Rowing Boat[1]	4.92	Camel[1]	5.88	Lion[1]	6.04
Aircraft Carrier[1]	3.20	Pheasant[8]	5.52	Helicopter[1]	6.12
Seahorse[8]	4.62	Oil Tanker[1]	5.48	Leopard[1]	6.06
Caravan[1]	4.52	Tractor[1]	5.32	Gorilla[8]	6.06
Submarine[1]	5.28	Snail[6]	5.30	Woodpecker[1]	6.00
Ostrich[8]	5.24	Crocodile[8]	5.90	Lorry[2]	6.00

Pictures sourced from: [1] https://pixabay.com [2] https://www.commons.wikimedia.org

[3] www.pngall.com [4] https://pngimg.com [5] https:// www.goodfreephotos.com

[6] www.freestockphotos.biz [7] Hatfield Image Test (Adlington et al., 2009) [8] The author

2.4 Design

In a narrative cover story, children were introduced to a forgetful and confused character, Peter the Panda, and asked to provide help with specific tasks that included:

2.4.1 ***Naming pictures:*** Children were asked to identify 48 objects by name, sampled from the typicality subset of the LOFTS semantic battery (Rogers et al., 2015). Answers were classified as correct if two or more adult participants had given the same name. Participant responses were recorded using a headset with a microphone that generated response times with millisecond accuracy. Response times were calculated as the time elapsed between picture onset and the onset of naming, for accurate responses only. The dependent measures for the naming task were accuracy and response latency to the onset of the correct object name.

2.4.2 ***Item (word-picture) matching:*** Children were required to match a probe word to a target image from an array of five images (compared to seven images available for adult participants in Rogers et al., 2015). Two variants of the word-picture matching task (proximal and distal conditions) were administered where the semantic distance between the target image and the distractor images in the array varied. Both conditions used identical sets of n = 48 probe words sampled from the typicality subset of the LOFTS semantic battery (Rogers et al., 2015). For the 48 trials in the proximal condition, the distractor images were closely related to the target image as members of the same category. For the 48 trials in the distal condition, distractors were less closely related as members of the animal kingdom. For example, a cheetah was presented with a shark, a kingfisher, an emu and a crocodile in the condition where semantic distance between the target and distractors was at its highest. The cheetah appeared as the target again with an elephant, a wildebeest, a moose and a skunk. Accurate identification of the target image was the dependent measure.

2.4.3 ***Sorting by category:*** Children were asked to sort 48 objects sampled from the typicality subset of the LOFTS semantic battery (Rogers et al., 2015). Two variants of the sorting task were generated that varied by the specificity of the category; a general category sort and a specific category sort. For the general category sort, children had to decide whether the objects belonged to one of three categories: animals, plants or vehicles. No plants were included in the set of pictures, however, the inclusion of this category enabled a 0.33 chance of successful guessing, thereby decreasing the chance of guessing correctly overall. For the specific sorting task children had to sort the animals and vehicles into the following categories: air, land and water. The two conditions of the sorting task were presented in separate test sessions to avoid repetition effects of including identical probe items in each task. All probe images were presented in a fixed random order that varied across the two tasks. The accuracy of responses was collated for all participants and the dependent variable was the number of correctly sorted objects.

2.4.4 ***Category matching:*** Children were required to match an image of a probe object to the image of a target object selected from an array, based on a common category between the probe and target objects. Forty-eight probe objects were sampled from the typicality subset of the LOFTS semantic battery (Rogers et al., 2015) and presented in a fixed random order. Two variants of the category matching task were administered with either proximal or distal distractors in the array. The semantic distance between the target image and the distractor images in the array varied by the specificity of the common category, similar to the design of the word-picture matching task. For example, in both the proximal and the distal conditions, participants had to match *tiger* with *lion*. However, in the proximal condition they needed to know that both of the items were from the shared specific category *'feline'*. This is because the distractors (*donkey, gorilla, wildebeest* and *skunk*) are closely related to the target in that they are all mammals. In the distal condition, it was only necessary for the child to know that tigers and lions are both mammals because the distractors (*crocodile, robin, salmon* and *jellyfish*) are not closely related to the target. The number of accurate selections of the target image was recorded.

2.4.5 ***Object-use (picture-picture) matching:*** This task involved matching photographs of 36 probe objects to photographs of the recipients of their functions. For example, children had to match a *pencil sharpener* to a *pencil* rather than *rubber, pen, ruler* or *crayon*. It was similar to the matching to recipient object-use test used by Corbett and colleagues (2009). The same number of probes was used as in the previous study, but different items were selected to be objects that children would be familiar with. In the current experiment, children had five items to choose from in comparison to four in the study by Corbett and colleagues (2009). Five items were used for consistency across the item matching and relational matching tasks. The probe objects were presented in a fixed random order. On each trial, children were required to select a matching target image from an array of five images, by adhering to the requirement to match by function. Two variants of the object-use matching task gave either proximal or distal distractors in the array. The distractors are either closely related to the targets or this relationship is distal. For example, participants had to match *scissors* with *sellotape* in both conditions. In the proximal condition, the distractors were: *glue, ruler, compass* and *paints* as compared to the distal condition where the distractors were: *basketball, shuttlecock, drum set* and *electric toothbrush*. The number of correct selections of the matching image was calculated.

For the probe objects and their recipients, photographic images of objects that children were expected to be familiar with were generated for this task. Thirty-six items (shown in Table 3) were sampled from ten different categories: ten kitchen utensils, two from beauty equipment, eight household tools, three from furniture, one personal accessory, three stationery items, two from bathroom equipment, five from sports equipment, one container and one object that makes a sound. Photographic images were collected from a variety of sources including the Hatfield Image

Test (Adlington et al., 2009), the Bank Of Standardized Stimuli (Brodeur, Dionne-Dostie, Montreuil & Lepage, 2010) and the Bank Of Standardized Stimuli (Brodeur, Guérard & Bouras, 2014), with other photographic images purchased using Shutterstock (an online source). Details of stimuli and their sources are listed in Table 3.

Table 3: List of photographic stimuli used in the object-use items

Pencil[1]	Comb[1]	Spoon[1]
Picnic Table[3]	Pencil Sharpener[1]	Hanger[1]
Match[1]	Fork[1]	Goal[4]
Eye Patch[3]	Badminton Racket[3]	Whistle[4]
Guitar Case[3]	Hammer[2]	Liquid Soap[1]
Fridge[3]	Office Chair[3]	Bookshelf[3]
Glass[1]	Nail Clippers[1]	Oven[3]
Watering Can[1]	Electric Toothbrush[3]	Tennis Racket[3]
Drumstick[3]	Sponge[1]	Muffin Tray[3]
Whisk[3]	Scissors[1]	Paintbrush[4]
Anchor[3]	Kettle[1]	Rollerblade[3]
Grater[3]	Knife[3]	Basketball Hoop[3]

Pictures sourced from: [1] The Bank Of Standardized Stimuli (Brodeur et al., 2010)

[2] Hatfield Image Test (Adlington et al., 2009)

[3] Bank Of Standardized Stimuli (Brodeur et al., 2014) [4] www.shutterstock.com

2.5 Procedure

All testing was carried out in areas of schools that were relatively quiet, as well as being overseen and overheard with ease. Participants sat at a child-sized table to complete the activities. The tasks were administered in a fixed random order within two separate sessions. All tasks were presented as computer games using DMDX (Forster & Forster, 2003) on a laptop computer. The order of the sessions was counterbalanced and each one lasted approximately 35 minutes.

2.5.1 Naming pictures: Children were presented with a character, Peter the Panda, and were asked to help him by reminding him of some names of objects he had forgotten. They were administered 48 coloured photographs of objects, one by one, in a fixed random order. Children were asked to say the names of the objects out loud and as clearly as they could. Participant responses were recorded using a headset with a microphone that generated response times with millisecond accuracy. Latencies were checked using CheckFiles to make sure they reflected the onset of naming rather than how long it took for the participant to make an irrelevant sound. The experimenter pressed the space bar to start a trial when the child was focused. They had three seconds to respond to each image. If a child corrected an original answer, the final response given was recorded.

2.5.2 Item (word-picture) matching: The two conditions of the word-picture matching task were presented in separate test sessions to avoid repetition effects of including identical probe words in each task. All probes were presented in a fixed random order that varied across the two tasks. Children were presented with a character, Peter the Panda, who was confused about words, and were asked to help him by selecting the picture that exactly matched a probe word. The experimenter started a trial by pressing the space bar when the child was paying attention. Over 3 seconds, children heard and saw a single word written onscreen. After this multimodal presentation of the probe word, they were required to choose the corresponding image by pointing to an array of five photographs, including the target image. All images remained on the screen until an item was chosen.

2.5.3 Sorting by category: Children were presented with a character, Peter the Panda, and told that he had forgotten where some items go. They were asked to help him out by sorting the objects into different boxes. When the participant was judged to be ready the experimenter started a trial by pressing the space bar. For the first four trials, a probe image of a single object was presented onscreen for 3 seconds, then the child would hear a question and see the same question written onscreen for 3 seconds. After this, the probe image was positioned above three pictures (with written category labels) that represented the three categories that they had to choose from. All images stayed onscreen until the child responded. For the subsequent 44 trials, the child simply saw the probe image as well as the three onscreen categories.

2.5.4 Category matching: The two conditions of the category matching task were presented in separate test sessions to avoid repetition effects of including identical probe pictures in each task. Children were presented with a character, Peter the Panda, and were asked to help him choose pictures that were 'like' one another. When the participant was focused the experimenter started a trial by pressing the space bar. First, the probe picture was displayed for 3 seconds. Then, participants saw an array of five images to choose from; one was the target image and the other four were distractors. Children were told that they should choose the picture that was most like the one they had just seen to help Peter. All images stayed on the screen until a selection was made.

2.5.5 Object-use (picture-picture) matching: The two conditions of the object-use matching task were presented in separate test sessions to avoid repetition effects of including identical probe pictures in each task. Children were introduced to a character, Peter the Panda, and told that he had forgotten how to use some objects. They were told to select the image that best indicated the use of the object. The experimenter started a trial by pressing the space bar when the child was judged to be ready. First, they were presented with a probe image for 3 seconds. Then, five images appeared and remained on the screen until the child responded.

2.6 Data Analysis

Data analysis was conducted using R (R Core Team, 2017). For all semantic measures, a multivariate modelling approach was applied. The use of a repeated measures ANOVA has been criticised because it is based on group level data and, therefore, ignores information that could be valuable at the level of the single item and individual participant (Wainwright, Leatherdale & Dubin, 2007). Therefore, accurate performance was analysed using generalised linear mixed models for binomially distributed outcomes (Williams, 1982). The following R packages were applied: 'lme4' (Bates, Maechler, Bolker & Walker, 2015) and 'car' (Fox & Weisberg, 2011) to generate the models and provide analysis of the fit to these data. 'LmerTest' (Kuznetsova, Brockhoff & Christensen, 2017) was used to obtain p-values. The majority of models used logit link functions and were robust to the non-normal distributions in these data i.e. Gaussian distributions were fitted that centred around the modes of the distributions with Laplace approximations (Wolfinger, 1993). The fixed and random effects estimates were calculated using log-odds. One exception was the naming task, where response latencies were analysed using linear mixed effects regression models fitted using the 'lme4' package. This approach was selected in preference to ANOVA because the latter is inappropriate for designs that include more than one random factor (Judd, Westfall & Kenny, 2017). In all models, participant and item were included as random factors. The R packages 'effects' and 'ggplot2' were implemented for graphical display and exploration of the effects.

3 Results

3.1 Object identification: Picture Naming accuracy and latency

The majority of children were able to accurately name at least 60% of these pictures of common objects. Figure 1 shows naming accuracy and latency for high, medium and low typicality items for each of the age groups. It suggests that children of all ages were better able to name the more typical objects. Quantile-quantile (Q-Q) plots (see Appendix C) were used to investigate the distribution of the accuracy data. The plots suggest that the data is not normally distributed. Shapiro-Wilk tests were also carried out. The accuracy scores for three age groups; 60-78-month-olds, $W = 0.76$, $p < 0.001$, 79-95-month-olds, $W = 0.82$, $p < 0.001$ and 96-130-month-olds, $W = 0.61$, $p < 0.001$, were all significantly non-normal.

To evaluate the accuracy of naming objects of varying typicality, by children of different age groups, a generalised linear mixed-effects model (GLMM) was generated to investigate age and typicality as fixed effects of interest. Since both fixed effects were coded as interval data, these two variables were centred before being entered into the model. This model also included participant and object as random effects that could also potentially influence the probability of successful object naming. A significant intercept indicated that the results did not occur by chance ($B = 1.22$, $SE = 0.23$, $Z = 5.22$, $p < 0.001$). Using a centred variable for typicality, there was a significant fixed effect ($B = 0.44$, $SE = 0.22$, $Z = 1.99$, $p < 0.05$) on naming accuracy. This means that the likelihood of naming a picture correctly depended on whether its typicality rating was high, medium or low. Using a centred variable for age group, there was also a significant fixed effect on naming accuracy ($B = 0.51$, $SE = 0.19$, $Z = 2.73$, $p = 0.01$), indicating that the likelihood of successful naming varied across age groups.

Figure 1a: Naming accuracy

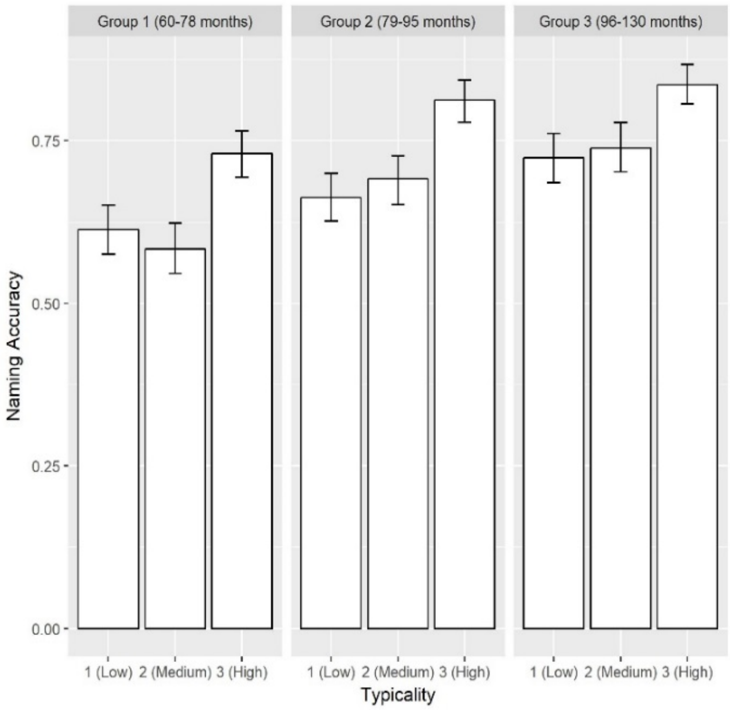
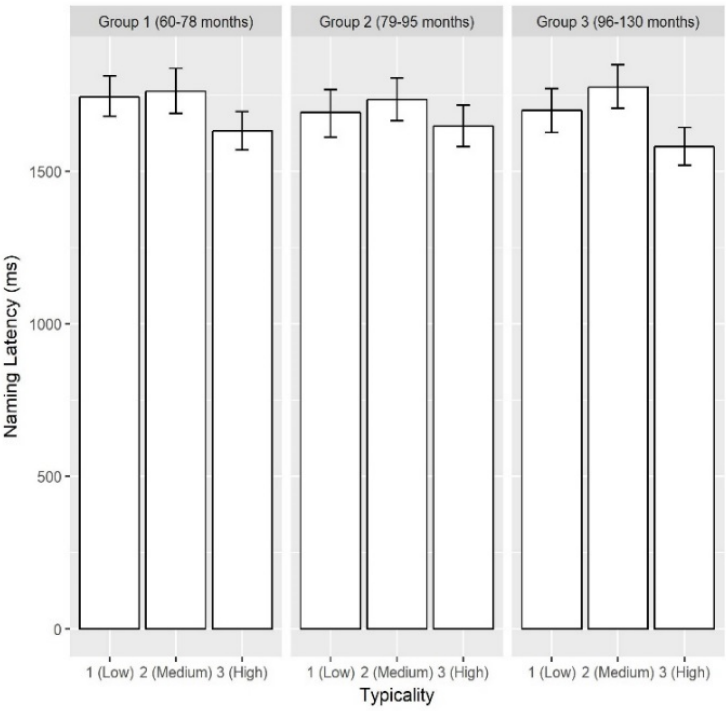


Figure 1b: Naming latency



Figures 1a and 1b. Mean accuracy (Fig. 1a as proportion correct) and latencies (Fig. 1b) for high, medium and low typicality objects as a function of age group. Error bars reflect 95% confidence intervals.

An ANOVA was used to compare the goodness of fit of this model with a baseline model that included random effects only, using Wald statistics estimated from the chi-square distribution. The ANOVA revealed that including the fixed effects of typicality and age significantly improved the fit of the model, $X^2(2) = 10.96, p < 0.01$. Therefore, age and typicality affected the likelihood of naming items accurately.

3.1.1 Investigating the interaction between age and typicality on picture naming accuracy

The experiential view predicted a main effect of age group only. By contrast, the CSC framework (Rogers et al., 2015) predicted a reverse typicality effect that interacted with age group. Figure 1a suggested a benefit for high typicality items and that children in all age groups were more accurate at naming highly typical pictures in a similar way across age groups. Following the prediction of an interaction between age and typicality, a further GLMM model of naming accuracy, using uncentred data, was generated to include an additional interaction term, as well as the fixed and random effects, consistent with the GLMM reported in 3.1. A significant intercept suggested that the results did not occur by chance ($B = 1.25$, $SE = 0.41$, $Z = 3.08$, $p < 0.01$). Consistent with the first fitted model, the fixed effect of typicality was significant; children were more likely to name highly typical images successfully compared to images with medium typicality ratings ($B = -0.92$, $SE = 0.45$, $Z = -2.05$, $p = 0.04$) and the effect of age was significant; the probability of naming items correctly was greater for the oldest children (96-130 months) as compared to the youngest age group (60-78 months; $B = 1.02$, $SE = 0.41$, $Z = 2.51$, $p = 0.01$). However, the interaction between age and typicality did not reach significance; more success with providing accurate names for pictures with high as compared to medium typicality ratings was similar for the oldest and the youngest children ($B = 0.12$, $SE = 0.25$, $Z = 0.50$, $p = 0.62$).

An ANOVA was used to compare the goodness of fit of the two models (with and without the interaction term) using Wald statistics estimated from the chi-square distribution. The ANOVA revealed that including the interaction between typicality and age did not significantly improve the fit of the model; the interaction was not significant, $X^2(4) = 3.46$, $p = 0.48$.

3.1.2 Identifying the locus of the effects for age group and typicality

To identify the origin of the effects of age group and typicality of the item to be named, the contrasts of estimated marginal means for the reduced model reported in 3.1 are displayed in Figure 2.

Figure 2: EMMs for naming accuracy

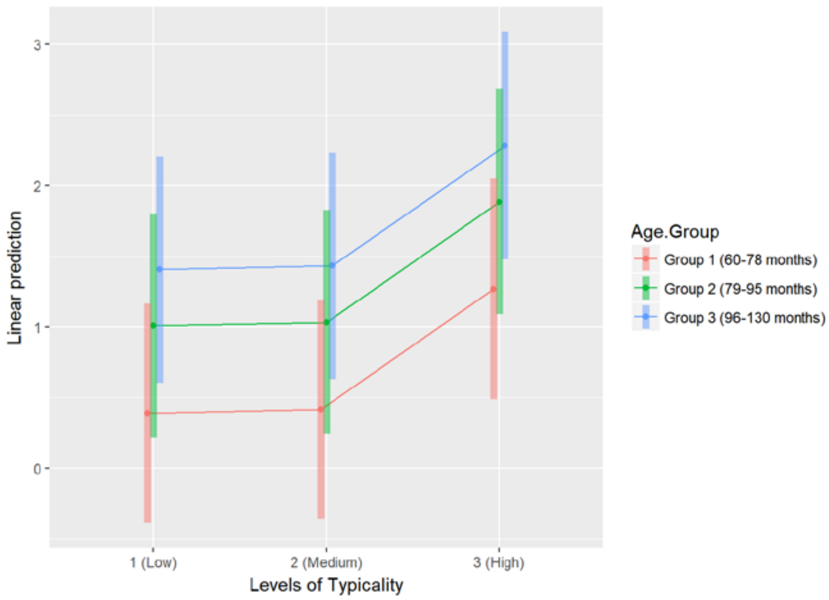


Figure 2: A plot showing EMMs for high, medium and low typicality objects as a function of age group. Error bars show 83% confidence intervals.

To determine the contrasts, the original variables, not centred variables, were used. The contrasts were used to determine the locus of any mean differences for typicality and age. Tukey adjustments were applied to p values as multiple comparisons were conducted. The first contrast for typicality showed that the mean difference between low and medium typicality did not reach significance ($B = -0.03$, $SE = 0.43$, $Z = -0.06$, $p = 1$), indicating that children were just as likely to name low as compared to medium typicality items correctly. The second contrast revealed that the difference between low and high typicality images was not significant ($B = -0.88$, $SE = 0.44$, $Z = -2.02$, $p = 0.11$). Children were equally likely to give the correct name if the picture had a low typicality rating as compared to a high one. Finally, children were just as likely to name medium as compared to high typicality items successfully ($B = 0.85$, $SE = 0.44$, $Z = -1.96$, $p = 0.12$).

The contrasts for age showed that the oldest children (96-130 months) were significantly more likely to show accurate naming than the youngest children (60-78 months; $B = -1.02$, $SE = 0.38$, $Z = -2.71$, $p = 0.02$). However, the age effect was not a linear trend. Additional contrasts compared the youngest group with the middle group (79-95 months: $B = -0.62$, $SE = 0.37$, $Z = -1.67$, $p = 0.22$) and the middle group with the oldest group ($B = -0.40$, $SE = 0.38$, $Z = -1.04$, $p = 0.55$) and found these age groups did not differ significantly in the likelihood of successful naming.

3.1.3 Investigating picture naming latencies

Since children showed high levels of accurate picture naming, their naming latencies were explored. Since response latencies were only collated for accurate trials, responses for incorrect trials were discarded. Overall, 29.2% of the reaction time data was discarded or unavailable, that ranged from 35.8 % in the 60-78-month-olds to 23.3% in the 96-130-month-olds. Unavailable data made up 5.6% of the reaction time data. This meant that 23.6% of the response latencies were removed due to incorrect responses, ranging from 27.9% in the 60-78-month-old group to 17.5% in the 96-130-month-old group.

Quantile-quantile plots (see Appendix D) were used to investigate the distribution of the latency data. The plots suggested that these data were not normally distributed. Shapiro-Wilk tests were conducted on the naming latencies for age group one (60-78 months), $W = 0.91$, $p < 0.001$, age group two (79-95 months), $W = 0.9$, $p < 0.001$ and age group three (96-130 months), $W = 0.9$, $p < 0.001$. These data were all significantly non-normal within each age group.

Figure 1b displays mean naming latencies for objects of varying typicality, as a function of age group and suggests that children in all groups were slightly faster at naming highly typical pictures. To confirm this pattern, a linear mixed-effects regression (LMER) model was generated to investigate age and typicality as fixed effects of interest. Additionally, participant and item were included to test random effects. As interval coded data, age group and typicality were centred before the main analysis was carried out. The fixed and random effects estimates were calculated using odds ratios. To reduce non-normality of the raw latency data, a log transformation was applied to these data before the analysis was carried out. A significant intercept indicated the results did not occur by chance ($B = 3.20$, $SE = 0.01$, $t(78.08) = 296.67$, $p < 0.001$). The effect of typicality as a centred variable on log-transformed naming latencies did not reach significance, $B = -0.02$, $SE = 0.01$, $t(47.45) = -1.42$, $p = 0.16$. This means that the likelihood of faster response times did not vary according to whether the image had a high, medium or low typicality rating. The latencies did not vary significantly with age as a centred variable, $B = -0.003$, $SE = 0.01$, $t(98.78) = -0.43$, $p = 0.67$. Age did not affect the probability of naming items quickly.

An ANOVA was used to compare the goodness of fit of this model with the baseline model that included random effects only, using Wald statistics estimated from the chi-square distribution. The ANOVA revealed that including the fixed effects of typicality and age did not significantly improve the fit of the model, $X^2(2) = 2.17$, $p = 0.34$. It was clear that age and typicality did not affect the likelihood of naming items more quickly.

3.1.4 Investigating the interaction between age and typicality on naming latencies

To explore the hypothesis that naming latencies of typical items varied systematically by age group, a further model of naming latency included an additional interaction term, along with the fixed effects of age and typicality as well as random effects of participant and item from the initial model. Centred variables of age and typicality were used in this model. A significant intercept suggested that the results did not occur by chance (B = 3.20, *SE* = 0.01, $t(78.09) = 296.68$, $p < 0.001$). As reported above, items were not likely to be named more quickly according to their typicality, $B = -0.02$, SE = 0.01, $t(47.50) = -1.43$, $p = 0.16$ or due to the age group, $B = -0.003$, SE = 0.01, $t(98.98) = -0.44$, $p = 0.66$. Since the interaction between age and typicality did not reach significance, $B = 0.0007$, SE = 0.004, $t(3501.00) = 0.19$, $p = 0.85$, typicality did not affect the likelihood of faster naming and this was true for all age groups. An ANOVA was used to compare the goodness of fit of the model that included a typicality * age interaction term with the reduced model that did not include the interaction term. The ANOVA revealed that including the interaction between typicality and age did not improve the fit of the first model, $X^2(1) = 0.03$, $p = 0.85$.

3.2 Object identification: Item (word-picture) matching accuracy

Consistent with their success at naming pictures, most children were equally able to accurately match spoken object names to their corresponding picture, with item matching accuracy scores greater than 75% regardless of the type of distractor used. Preliminary exploration of these data for all children indicated that scores were not normally distributed for the low, medium and high typicality items when presented with proximal distractors (Shapiro-Wilk $W = 0.59$, $p < 0.001$) or distal distractors (Shapiro-Wilk $W = 0.53$, $p < 0.001$- see also Appendix E for Q-Q plots).

Figure 3: Item matching accuracy

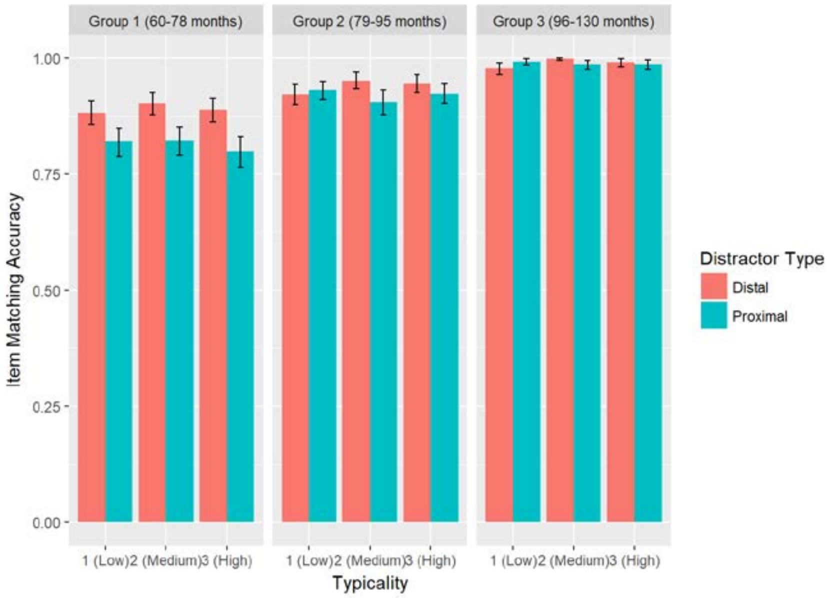


Figure 3. Mean item matching accuracy (proportion correct) for high, medium and low typicality pictures as a function of age group. Error bars show 95% confidence intervals.

As shown in Figure 3, children in the older age group (>95 months) performed approximately at ceiling so their item matching scores were not analysed further. Figure 3 further suggested that children in the younger age group (<79 months) showed more variability when matching items using different types of distractor compared to more stable performance in older children aged 79-95 months.

3.2.1 Investigating the interaction between age and typicality on word-picture matching in young children

A GLMM was fitted to these item matching data to determine whether the log odds of successful matching varied systematically by the inclusion of an interaction term for age* typicality of items, whilst accounting for the fixed effects of age, typicality and distractor type along with random effects of participant and item. The data for typicality was centred for this analysis. A significant intercept suggested that the results did not occur by chance (B = 2.96, *SE* = 0.27, Z = 11.11, p <

0.001). The effect of age group reached significance (B = 0.98, SE = 0.29, Z = 3.34, p < 0.001) and so did the effect of distractor type (B = -0.74, SE = 0.09, Z = -7.99, p < 0.001) but the effect of typicality did not (B = -0.01, SE = 0.21, Z = -0.06, p = 0.95). The age*typicality interaction did not reach significance either (B = 0.20, SE = 0.12, Z = 1.74, p = 0.08), implying that the older children of 79- 95 months were more likely than younger children to successfully match the items regardless of the typicality of the picture.

3.2.2 Investigating the influence of task distractors on word-picture matching in young children

A hypothesis motivated by the CSC framework was that the type of task distractor could influence successful matching as a function of the typicality of the item. To address this, two further GLMM analyses were carried out to directly assess the influence of distractors in relation to age and typicality (reported in 3.2.3).

A GLMM was fitted to item matching data to estimate whether the log odds of successful matching varied systematically by the inclusion of an interaction term for distractor type * age, whilst accounting for the fixed effects of distractor type, age and typicality as well as random effects of participant and item. Typicality was centred before it was entered into the model. A significant intercept indicated that results did not occur by chance (B = 3.08, SE = 0.27, Z = 11.36, p < 0.001). Consistent with the first model, this GLMM confirmed that the effect of typicality did not reach significance (B = 0.05, SE = 0.21, Z = 0.25, p = 0.80). There was a significant effect of age group (B = 0.64, SE = 0.31, Z = 2.05, p = 0.04) in that older children (79-95 months) were more accurate at item matching than younger children. The effect of distractor type was also significant (B = -0.94, SE = 0.12, Z = -8.00, p < 0.001) indicating successful matching was more likely when the distractors were distal, rather than proximal. In addition, the interaction between distractor type and age was significant (B = 0.54, SE = 0.19, Z = 2.85, p < 0.01). Contrasts of the estimated mean scores revealed that, when faced with proximal distractors, the probability of accuracy was greater for older, as compared to younger, children ($B = -1.19$, $SE = 0.31$, $Z = -3.89$, $p < 0.001$). However, when distractors were distal, the likelihood of successful matching did not vary by age group ($B = -0.64$, $SE = 0.31$, $Z = -2.05$, $p = 0.17$). All contrasts were subject to Tukey adjustments for multiple comparisons to avoid type 1 error in reporting mean differences.

An ANOVA was used to compare the goodness of fit of this model that included distractor type * age to a reduced model that did not include the interaction term. The ANOVA revealed that including the interaction between distractor type and age improved the fit of the model, $X^2(1) = 7.67$, $p = 0.01$.

We also explored the age by typicality by distractor type interaction but this is not reported as we are not confident that our study had sufficient power to detect this interaction.

3.2.3 Identifying the locus of the effects for typicality by distractor

A further GLMM was fitted to item matching data to evaluate the influence of distractor type on matching accuracy for items that varied by typicality. This model included a typicality*distractor type interaction term that estimated the log odds of successful matching, whilst accounting for fixed effects of age group, distractor type and typicality, as well as the random effects of participant and item, as reported for an earlier model in 3.2.1. A significant intercept suggested that results did not occur by chance ($B = 2.97$, $SE = 0.27$, $Z = 11.14$, $p < 0.001$). Consistent with the earlier model, using a centred variable for typicality the effect did not reach significance ($B = 0.22$, $SE = 0.22$, $Z = 1.01$, $p = 0.731$) but the effects of age ($B = 0.97$, $SE = 0.29$, $Z = 3.29$, $p < 0.001$) and distractor type were significant ($B = 0.76$, $SE = 0.09$, $Z = 8.12$, $p < 0.001$). The interaction between typicality and distractor type was also significant ($B = 0.28$, $SE = 0.11$, $Z = 2.48$, $p = 0.01$), implying that the effects of distractor varied by the typicality of the item.

An ANOVA was used to compare the goodness of fit of this interaction model with a reduced model that did not include any interaction term. The ANOVA revealed that including the interaction between distractor type and typicality improved the fit of the model, $X^2(1) = 5.92$, $p =$

0.01. To identify the origin of the effects of distractor, according to the typicality of the item to be matched, the contrasts of estimated marginal means are displayed in Figure 4.

Figure 4: EMMs for item matching accuracy

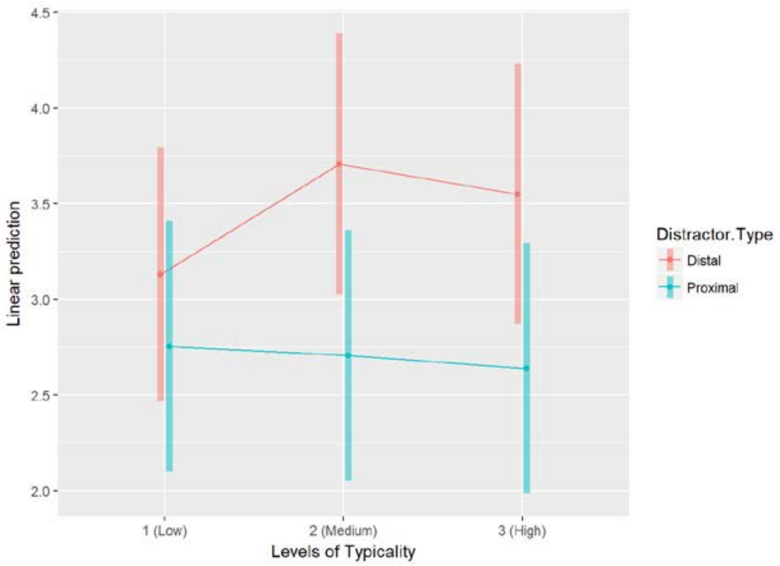


Figure 4: A plot showing estimated mean matching scores for high, medium and low typicality items with proximal and distal distractors. Results are averaged over the levels of: age group. Error bars show 83% confidence intervals.

As shown in Figure 4, item matching scores with the proximal distractors appeared stable for all levels of typicality. By contrast, more variability in item matching according to typicality was observed with distal distractors.

Contrasts of these estimated mean scores revealed that the probability of accurate matching was greater with distal, as compared to proximal, distractors when items were both of high (B = -0.91, SE = 0.17, Z = -5.45, p < 0.001) and of medium typicality (B = -1.00, SE = 0.17, Z = -6.05, p < 0.001). However, for less typical items, matching did not vary with the nature of the distractors used (B = -0.38, SE = 0.15, Z = -2.5, p = 0.13). All contrasts were subject to Tukey adjustments for multiple comparisons to avoid Type 1 error in reporting mean differences.

3.3 Object discrimination: Picture sorting accuracy

Children of all ages were able to accurately sort pictures by their general and specific categories, as shown in Figure 5. Preliminary exploration of these data for all children indicated that scores were not normally distributed for the low, medium and high typicality items when sorting into general (Shapiro-Wilk $W = 0.39$, $p < 0.001$) or specific categories (Shapiro-Wilk $W = 0.66$, $p < 0.001$ - see also Appendix F for Q-Q plots).

On the general sorting task, performance approximated ceiling for children above the age of 78 months, as shown in Figure 5a. For younger children (<79 months), sorting accuracy was high on both general and specific sorting tasks, as shown in Figures 5a and 5b respectively. Therefore, the data for both sorting tasks were analysed for the youngest group only.

Figure 5a: General sorting accuracy

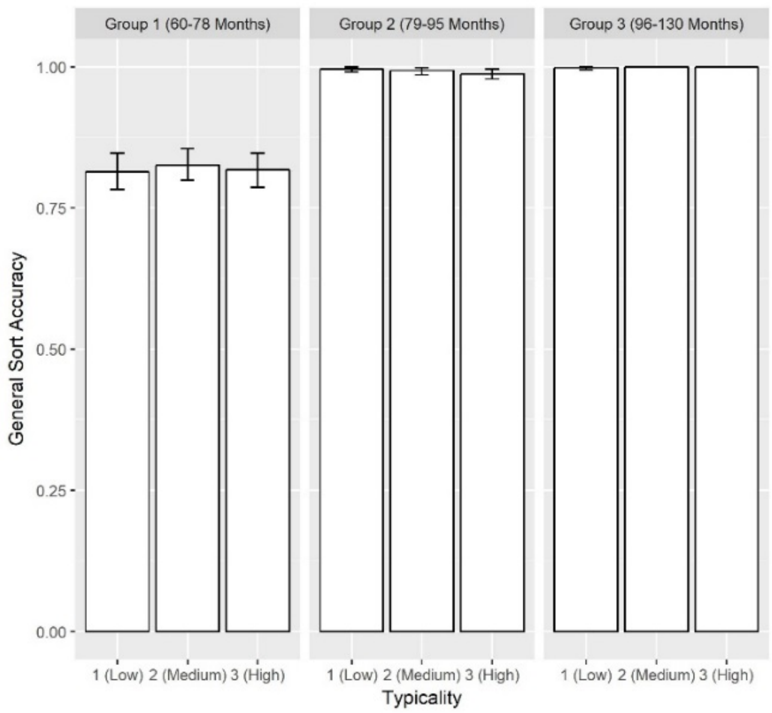
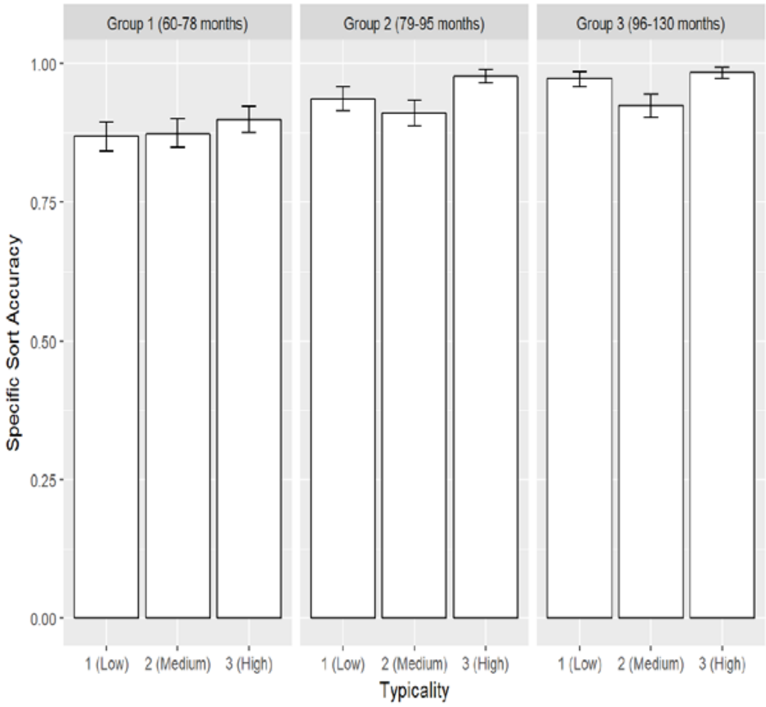


Figure 5b: Specific sorting accuracy



Figures 5a and 5b. Mean sorting accuracy (as proportion correct) for high, medium and low typicality items using a generic sortal (Figure 5a) and a specific sortal (Figure 5b) by age group. Error bars show 95% confidence intervals.

3.3.1 Investigating general and specific sorting in the youngest age group (60-78 months)

Visual inspection of the youngest children in Figures 5a and 5b suggested that the use of a specific sortal term elicited a small improvement in their sorting ability. To explore this, a GLMM model was fitted to the data from both sortal tasks to estimate the log odds of successful sorting according to the specificity of the sortal term. Typicality was also included in the model as a centred variable as well as participant and item as random effects that could affect the probability of successful sorting. A significant intercept indicated that results did not occur by chance (B = 2.20, SE = 0.27, Z = 8.05, p < 0.001). The effect of specificity was significant (B = 0.68, SE = 0.11, Z = 6.13, p < 0.001) such that the probability of accurate sorting was greater when category sortal was specific rather than general, as shown in Figure 6. Children were just as likely to sort low, medium and high typicality pictures correctly ($B = 0.09$, $SE = 0.13$, $Z = 0.68$, $p = 0.50$).

Figure 6: EMMs for general and specific sorting accuracy (60-78 months)

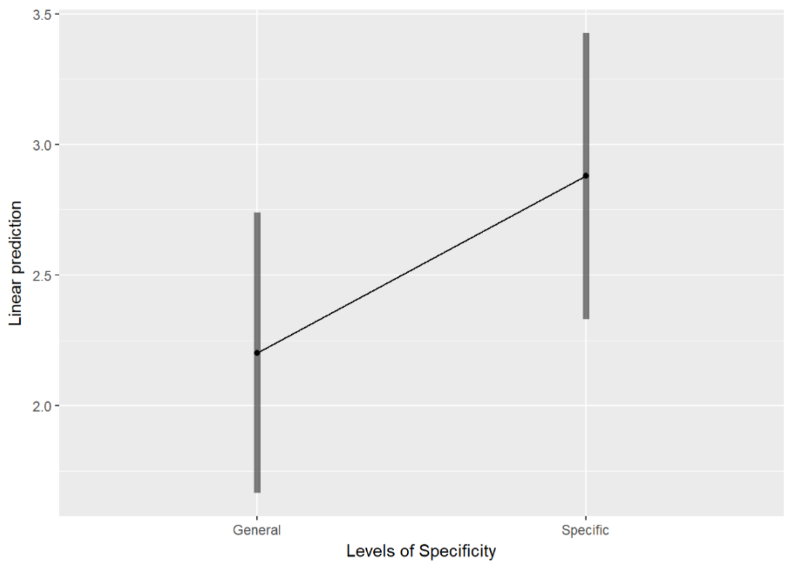


Figure 6: Estimated marginal mean accuracy scores for general and specific sortals in young children aged 60-78 months. Error bars show 83% confidence intervals.

An ANOVA was used to compare the goodness of fit of this model with the baseline model that only included the random effects of participant and item. By including the fixed effects of specificity and typicality, the fit of the model improved for the youngest children only, $X^2(1) = 37.25$, $p < 0.001$. Children under 79 months were more likely to be accurate at sorting pictures into categories using a specific, as compared to a general, sortal.

The CSC framework motivated a further hypothesis that the specificity of the sortal term could influence accuracy as a function of the typicality of the item. To address this, a GLMM analysis was carried out that included a typicality*specificity interaction that aimed to estimate the log odds of successful sorting, whilst accounting for fixed effects of typicality and specificity in addition to random effects of participant and item that could influence successful sorting. A significant intercept suggested that results did not occur by chance ($B = 2.20$, $SE = 0.27$, $Z = 8.04$, $p < 0.001$). Using a centred variable for typicality, the effect did not influence the probability of accurate sorting overall ($B = 0.02$, $SE = 0.15$, $Z = 0.11$, $p = 0.91$). The specificity of the sorting categories did influence likelihood of accuracy ($B = 0.68$, $SE = 0.11$, $Z = 6.17$, $p < 0.001$). The interaction between specificity and typicality did not reach significance ($B = 0.17$, $SE = 0.13$, $Z = 1.3$, $p = 0.19$). Therefore, the effect of specificity was not dependent on the typicality of the pictures, when discriminating between objects from different classes.

An ANOVA was used to compare the goodness of fit of the specificity * typicality model with a model that did not include the interaction (this model considered fixed effects of specificity and typicality as well as random effects of item and participant). The ANOVA revealed that including the interaction between specificity and typicality did not improve the fit of the model, $X^2(1) = 1.63$, $p = 0.2$. Therefore, we conclude that the effect of specificity was not dependent on the typicality of the images.

3.3.2 Investigating age and typicality effects on specific sorting only (all groups)

As shown in Figure 5, sorting accuracy was more variable for older children when using a specific sortal. A GLMM was fitted to data from children of all age groups to estimate the log odds of successful sorting according to age group and typicality. Typicality and age were entered as centred variables. Participant and item were included as random effects that could affect the probability of successful sorting. A significant intercept indicated that the results did not occur by chance ($B = 3.14$, $SE = 0.32$, $Z = 9.86$, $p < 0.001$). Overall, there was an effect of typicality ($B = 0.27$, $SE = 0.09$, $Z = 3.05$, $p < 0.01$), suggesting that the likelihood of accuracy varied according to whether typicality was low, medium or high. Age did not affect the probability of sorting pictures accurately ($B = 0.42$, $SE = 0.23$, $Z = 1.80$, $p = 0.07$). The oldest (96-130 months) and middle (79-95 months) age groups were no more accurate than the youngest children (60-78 months).

To identify the origin of the effects of typicality of the item to be sorted, the contrasts of estimated marginal means from the fitted model are displayed in Figure 7.

Figure 7: EMMs for specific sorting accuracy

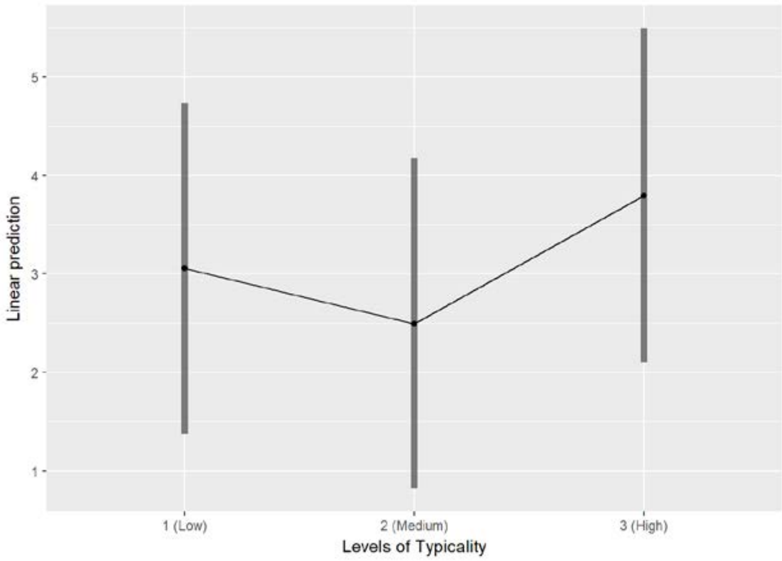


Figure 7 displays estimated marginal means for low, medium and high typicality pictures. Error bars show 83% confidence intervals.

The original typicality variable (rather than the centred one) was entered for direct comparisons of estimated marginal means at all levels of typicality. Children were more likely to sort low (B = 0.56, SE = 0.17, Z = 3.37, p < 0.01) and high (B = -1.30, SE = 0.20, Z = -6.37, p < 0.001) typicality images accurately as compared to medium ones, as shown in Figure 7. The likelihood of accuracy was also greater for high as compared to low typicality items (B = -0.74, SE = 0.21, Z = -3.46, p < 0.01).

An ANOVA was used to compare the goodness of fit of this model with the baseline model (that only included the random effects of participant and item). The ANOVA revealed that including the fixed effects of age group and typicality improved the fit of the model, $X^2(3) = 11.80, p = 0.01$.

Since the controlled cognition approach to conceptual development predicted that typicality would have a greater effect on sorting for the youngest children, a further GLMM was fitted to include an age group*typicality interaction. This model aimed to estimate the log odds of successful sorting, whilst accounting for fixed effects of age group and typicality as well as the random effects of participant and item. Typicality and age group were centred (as in the previous model). A significant intercept suggested that results did not occur by chance ($B = 3.13$, $SE = 0.32$, $Z = 9.82$, $p < 0.001$). The main effect of age on accuracy was not significant ($B = 0.41$, $SE = 0.24$, $Z = 1.75$, $p = 0.08$), indicating that the likelihood of accurate sorting was not significantly greater for the two older age groups as compared to the youngest children. As shown in Figure 7 and consistent with the previous model, the main effect of typicality on accuracy was significant ($B = 0.29$, $SE = 0.10$, $Z = 3.01$, $p < 0.01$). The typicality effect did not vary by age group ($B = -0.07$, $SE = 0.14$, $Z = -0.51$, $p = 0.61$).

An ANOVA was used to compare the goodness of fit of the age * typicality model with a reduced model that did not include the interaction term. Since the inclusion of the interaction between age and typicality did not improve the fit of the model, $X^2(2) = 1.52$, $p = 0.47$, it was clear that the probability of accurate sorting was affected by typicality and that age group did not alter this effect for specific sorting.

3.4 Object discrimination: Category matching (picture-picture) accuracy

All children were able to accurately match at least 75% of the probes to members of the same category when distractors were distal. Preliminary exploration of these data for all children indicated that scores were not normally distributed for the low, medium and high typicality items when presented with proximal distractors (Shapiro-Wilk $W = 0.90$, $p < 0.001$) or distally related distractors (Shapiro-Wilk $W = 0.83$, $p < 0.001$ - see also Appendix G for Q-Q plots).

Figure 8: Category matching accuracy

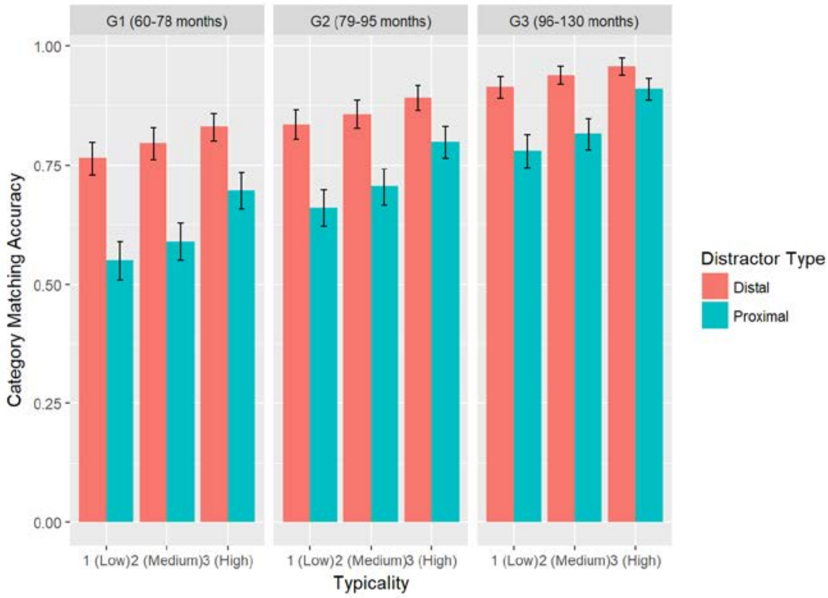


Figure 8. Mean category matching accuracy (proportion correct) for high, medium and low typicality pictures by age group. Error bars reflect 95% confidence intervals.

Figure 8 displays the category matching accuracy for children of all ages and indicates that children were better able to match pictures by category when distractors were distally related to the target object. This pattern will be explored further in the next section.

3.4.1 Investigating the interaction between age and typicality on category matching accuracy

One prediction was that older children would have more enriched knowledge about atypical items than younger children. A GLMM was fitted to these category matching data to estimate whether the log odds of successful matching varied systematically by the inclusion of an interaction term for age* typicality of items, whilst accounting for the fixed effects of age, typicality and distractor type, in addition to random effects of participant and item. Age and typicality were centred for this model. A significant intercept suggested that results did not occur by chance (B = 2.28, SE = 0.13, Z = 17.56, $p < 0.001$). The effect of age was significant ($B = 0.69$, $SE = 0.09$, $Z = 7.44$, $p <$

0.001), indicating that the likelihood of accurate matching depended on age. Likelihood of accurate matching also depended on typicality (B = 0.33, SE = 0.13, Z = 2.61, p = 0.01). The probability of accurate matching was greater when distractors were distal as compared to proximal (B = -1.11, SE = 0.06, Z = -19.31, $p < 0.001$). However, the interaction between typicality and age did not reach significance (B = 0.06, SE = 0.04, Z = 1.27, p = 0.21), implying that typicality effects did not vary systematically by age group.

3.4.2 Investigating the interaction between typicality and distractor on category matching accuracy

A prediction from the CSC framework was that the type of task distractor could influence successful matching as a function of the typicality of the item being matched. To address this, a second GLMM with centred age and typicality variables was fitted to category matching data. This model included a typicality*distractor type interaction term that aimed to estimate the log odds of successful matching, whilst accounting for fixed effects of age group, distractor type and typicality, as well as the random effects of participant and item. A significant intercept indicated that results did not occur by chance (B = 2.27, SE = 0.13, Z = 17.37, $p < 0.001$). Similar to the findings of the earlier model, age (B = 0.69, SE = 0.09, Z = 7.38, $p < 0.001$) and typicality (B = 0.39, SE = 0.13, Z = 3.03, $p < 0.01$) affected the likelihood of successful matching, and target pictures were more likely to be matched successfully when distractors were distal rather than proximal (B = -1.09, SE = 0.06, Z = -19.01, $p < 0.001$). However, the likely effect of matching typical items varied according to the distractor type (B = 0.20, SE = 0.07, Z = 2.87, $p < 0.01$).

To identify the pattern of age effects, estimated marginal means were generated. To adjust for multiple comparisons, Tukey adjustments were applied to p values. Overall, the oldest age group (96-130 months) scored better than a younger age group (79-95 months; B = -0.79, SE = 0.19, Z = -4.24, $p < 0.001$). The youngest age group (60-78 months) scored worse than older children (79-95 months; B = -0.58, SE = 0.18, Z = -3.25, $p < 0.01$).

An ANOVA was used to compare the goodness of fit of this model (that included the interaction term) with a reduced model that did not include the interaction. The ANOVA revealed that including the interaction between distractor type and typicality significantly improved the fit of the model, $X^2(1) = 8.08$, $p < 0.01$.

3.4.3 Identifying the locus of the effects for typicality by distractor

To identify the origin of the effects of distractor according to the typicality of the item to be matched by category, the contrasts of estimated marginal means for the interaction from the fitted model reported are displayed in Figure 9.

Figure 9: EMMs for category matching accuracy

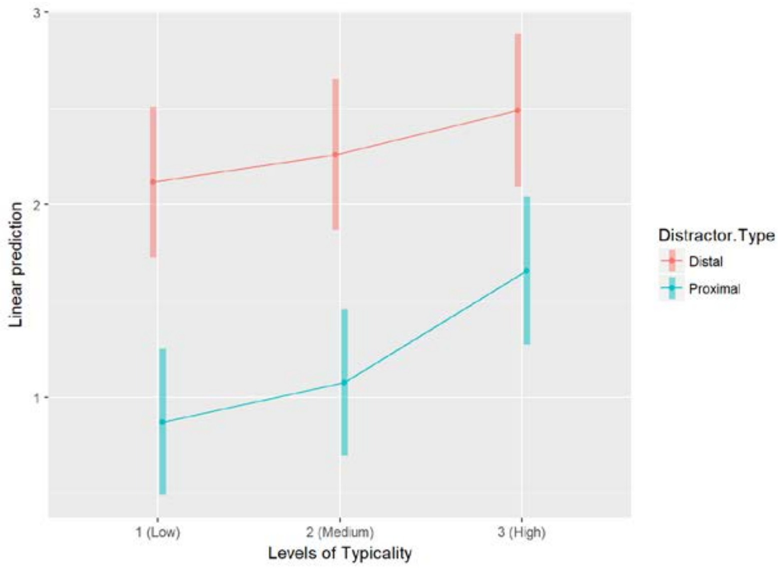


Figure 9 displays estimated marginal means for high, medium and low typicality items with proximal and distal distractors for each age group. Error bars show 83% confidence intervals.

None of the contrasts for typicality reached full significance; children were marginally more likely to match high as compared to low typicality pictures correctly ($B = -0.58$, $SE = 0.25$, $Z = -2.3$, $p = 0.06$) and no mean differences were observed for medium typicality as compared to low ($B = -$

0.17, $SE = 0.25$, $Z = -0.70$, $p = 0.77$) or high ($B = -0.40$, $SE = 0.25$, Z = -1.61, $p = 0.24$) typicality. However, these effects were qualified by typicality effects that varied with the distractor type. When comparing the high and low typicality images, typicality effects were evident when the distractors were proximal ($B = -0.79$, $SE = 0.26$, $Z = -3.06$, $p = 0.03$). However, when distractors were distal, there was no effect of typicality on the probability of accurate category matching ($B = -0.37$, $SE = 0.27$, $Z = -1.4$, $p = 0.73$).

3.4.4 Investigating the interaction between distractor type and age on category matching accuracy

A final consideration was whether children in different age groups were less likely to match pictures successfully, according to the type of distractor. As shown in Figure 8, category matching accuracy appeared to improve systematically with age. However, it tentatively appeared that children in the youngest age group (60-78 months) were less likely to match successfully in the context of proximal distractors. Using centred variables for age and typicality, a GLMM was fitted to these data, to include an age*distractor type interaction. The interaction term aimed to estimate the log odds of successful category matching according to age across distractor type, whilst accounting for fixed effects of typicality, distractor type and age, as well as the random effects of participant and item. A significant intercept suggested that results did not occur by chance (B = 2.28, $SE = 0.13$, Z = 17.47, $p < 0.001$). Consistent with the fitted model that explored the age*typicality interaction in 3.4.1, age affected the likelihood of matching images correctly ($B = 0.67$, $SE = 0.10$, $Z = 6.96$, $p < 0.001$), typicality also had an effect ($B = 0.32$, $SE = 0.13$, $Z = 2.51$, $p = 0.01$) and successful matches were more likely when distractors were distal, rather than proximal ($B = -1.11$, $SE = 0.06$, $Z = -18.73$, $p < 0.001$). However, the interaction between age and distractor type did not reach significance ($B = 0.04$, $SE = 0.07$, $Z = 0.51$, $p = 0.61$), suggesting that all age groups behaved uniformly according to distractor type.

3.5 Object-use matching accuracy

The object-use matching task extended the findings of the category matching test, reported in section 3.4, by varying the type of semantic relation and using a distinct set of items. For this task, the majority of children were able to accurately match at least 75% of the objects to the recipient of the object function. Preliminary exploration of these data for all children indicated that scores were not normally distributed when presented with proximal distractors (Shapiro-Wilk $W = 0.85$, $p < 0.001$) or distal distractors (Shapiro-Wilk $W = 0.83$, $p < 0.001$ - see also Appendix H for Q-Q plots).

Figure 10: Object-use accuracy

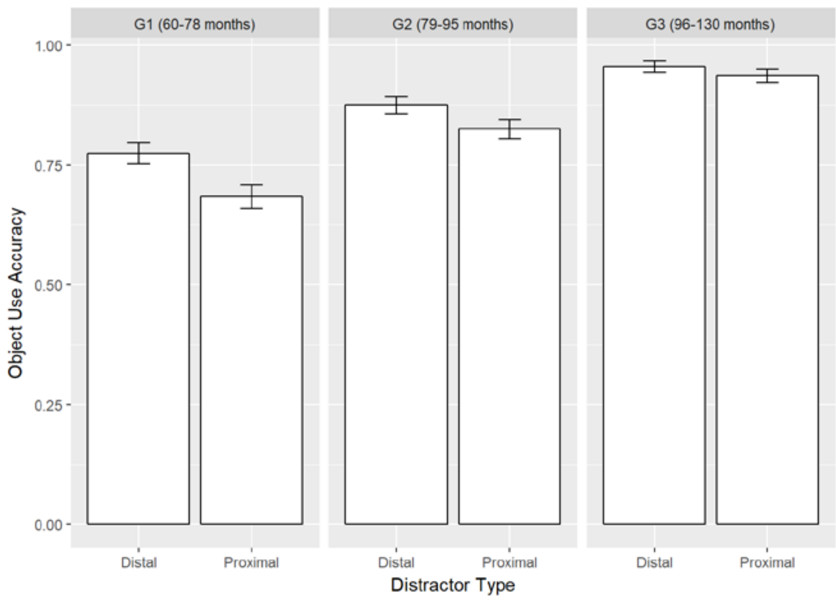


Figure 10. Mean object-use matching accuracy (as proportion correct) with proximal and distal distractors by age group. Error bars show 95% confidence intervals.

Figure 10 displays mean object-use matching accuracy for each of the age groups, by the relation of the distractors to the target object. It suggests that children of all ages were better able to match items when distractors were distally related to the target object, rather than proximally related.

To explore the influence of distractors and age on the likelihood of accurate object-function matching, a GLMM was fitted to these data. The model sought to estimate the log odds of successful matching according to age group and distractor type, with participant and item included as random effects. A significant intercept indicated that results did not occur by chance ($B = 2.45$, $SE = 0.15$, $Z = 16.10$, $p < 0.001$). Using a centred variable for age group there was a significant effect on the likelihood of accurate matching ($B = 1.03$, $SE = 0.12$, $Z = 8.8$, $p < 0.001$). Accurate matching was more likely when the distractors were distal, rather than proximally-related to the target object ($B = -0.50$, $SE = 0.07$, $Z = -7.2$, $p < 0.001$). For example, the probability of making a successful match between *pencil sharpener* and *pencil* was greater when the distractors were *guitar, keyboard, goal* and *wardrobe,* as opposed to *rubber, pen, ruler* and *crayon.*

An ANOVA compared the goodness of fit of this model to a baseline model that included only the random effects, using Wald statistics estimated from the chi-square distribution. The fixed effects of distractor type and age significantly improved the fit of the model to these matching data, $X^2(2) = 111.31$, $p < 0.001$.

3.5.1 Investigating the interaction between age and distractor type on matching accuracy

Since the objects were highly familiar to children, the experiential hypothesis only predicts that children of this age range will be able to match images of these familiar objects to their recipients, based on their experience of the objects. By contrast, the controlled cognition hypothesis predicts an influence of distractor type that would be greater for the youngest children.

A further GLMM was fitted to these object-use data. This model used centred data to estimate whether the log odds of successful matching varied systematically by the inclusion of an interaction term for age* distractor type, whilst accounting for the fixed effects of age and distractor type in addition to random effects of participant and item. A significant intercept suggested that results did not occur by chance ($B = 2.43$, $SE = 0.15$, $Z = 15.87$, $p < 0.001$). Consistent with the previous model, likelihood of successful matching was influenced by age ($B = 0.99$, $SE = 0.13$, $Z = 7.62$, p

< 0.001) and semantic relationship of the distractors to the target object (B = -0.47, SE = 0.08, Z = -6.13, p < 0.001). The interaction between age and distractor type did not reach significance (B = 0.08, SE = 0.09, Z = 0.88, p = 0.38), implying that the likelihood of successful matching was greater for distally-related distractors, regardless of age group.

An ANOVA was used to compare the goodness of fit of this model with a reduced model that did not include the interaction term. Including the interaction term between distractor type and age did not improve the fit of the model, $X^2(1) = 0.74$, $p = 0.39$.

3.5.2 Identifying the locus of the effects for age group

To identify the origin of the effect of age group, the contrasts of estimated marginal means for the reduced model reported in 3.5 are displayed in Figure 11.

Figure 11: EMMs for object-use accuracy

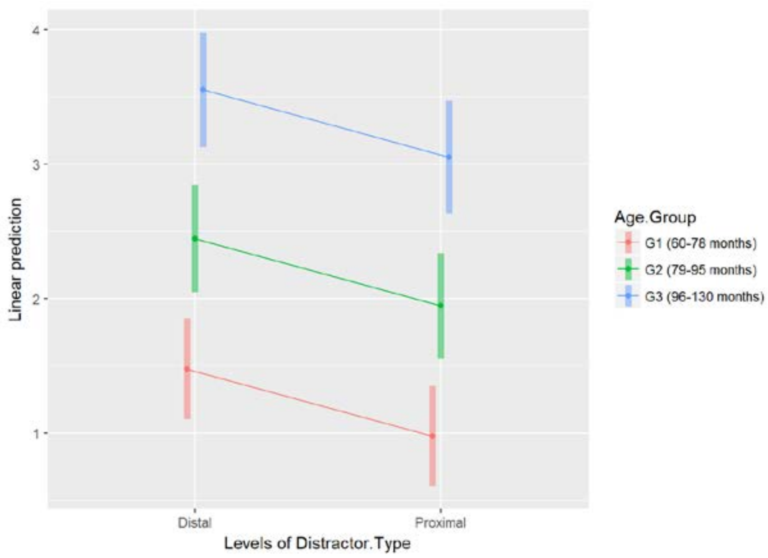


Figure 11. Estimated marginal means for the likelihood of successful object-use matching for each age group, as a function of distal and proximal relations of the distractors to the target object. Error bars show 83% confidence intervals.

Tukey adjustments were applied to p values as multiple comparisons were conducted. The contrasts for age group revealed that successful matching was more likely for the oldest age group (96-130 months) compared to the middle age group (79-95 months; $B = -1.11$, $SE = 0.24$, $Z = -4.54$, $p < 0.001$) and less likely for the youngest age group (60-78 months) compared to the middle age group (79-95 months; $B = -0.97$, $SE = 0.22$, $Z = -4.33$, $p < 0.001$).

4 Discussion

4.1 Overview of findings

For all semantic measures, accuracy levels were mostly high for the children in this study. This suggests that the items were known to the participants. Children gave the correct response for at least 60% of the trials in most tasks, with the exception for category matching when the distractors were proximally-related to the targets. Under these conditions, matching accuracy dropped to approximately 50%. Reduced accuracy cannot be due to the objects being unknown to the participants as identical items were used in the category matching task as in naming, sorting and item-matching. Overall, the tasks suggest that at least 60% of the probe items were known. In other words, children had accumulated a high level of knowledge about the concepts probed in the tasks, but tasks varied by the demands for controlled cognition.

4.2 Evaluating findings from the picture naming task

Successful picture naming requires identification of an object name from amongst other competing names through access to its visual features. In this way, the task requires available concept knowledge and efficient selection of the correct name from competing alternatives for correct object identification i.e. the task is high in selection demands.

4.2.1 The experiential hypothesis and picture naming

The experiential view suggests that older children will be faster and more accurate at naming than younger children as they have accumulated more conceptual knowledge due to experience with a wider variety of items. The present findings suggested a naming advantage only for the oldest children in the age-range. This could provide partial support for an experiential view in that naming increases with age due to the accumulated knowledge of the older children. However, performance was not consistent across different tasks with identical items. All age groups performed at a superior level on other semantic tasks that involved identical concepts. This suggests that factors other than the amount of accumulated knowledge, also influenced performance by age. The relatively poor naming ability of the younger children was not consistent with a failure to 'know' the words. Rather, it could also reflect a relative difficulty with retrieving the words in a spoken form.

The experiential view could elicit a typicality advantage, if typicality is correlated with the age-of-acquisition of items. For example, if highly typical items are experienced earlier in childhood (Schröder et al., 2012) then they would be at the heart of the semantic system (Brysbaert & Ghyselinck, 2006). However, since the typicality sets were matched for familiarity, graded differences by typicality were not predicted for naming accuracy. Children were able to accurately name a good number of the low typicality pictures. Despite the matching by familiarity, there was a graded advantage for naming more typical images in children across the age-range. So, the defining characteristics that are more typical of a category appeared to provide stronger cues for retrieving the more typical items. Since typicality did not interact with age, the effects could be explained by an experiential account if knowledge of typical items is accumulated before that of atypical ones. The developmental characteristics suggest that structure of the semantic network supports richer category clusters for typical objects that provides a stable benefit for identifying typical objects across childhood.

4.2.2 The CSC framework and picture naming

A focus of the CSC framework (Rogers et al., 2015) is the nature of task requirements that engage control during retrieval processes, as well as the accumulated knowledge in a semantic network. For naming, the task places high demands on the selection of the correct candidate from multiple candidates that share features and semantic relations, as well as the lexical aspects of word production. In this way, the comparative difficulty of the naming task suggests that the retrieval demands of selection are significant in the naming task. By contrast, selection demands are not as great for sorting or item matching. This is because, in these tasks, children had a limited number of options to choose from.

The CSC framework makes strong predictions regarding the direction of typicality effects i.e. that a rich accumulation of shared semantic features and relations for typical objects will bring additional retrieval demands for the selection of the target item. This is because typical pictures bring to mind objects from the same category and this activity increases competition for selecting the correct target. For this reason, a reverse gradation of the typicality effect was predicted for children over the age of 78 months, once they can invoke controlled cognition to co-ordinate a richer semantic network. However, the findings did not support this prediction; the children of all ages were consistently more accurate at naming the highly typical images. Although the observed pattern of typicality effects did not support the CSC framework prediction, it is possible that the oldest children were more able to recruit controlled cognition to overcome the retrieval demands.

4.3 Evaluating findings from the item matching task

The item matching task measured the ability to identify a single object concept by matching to its label across different modalities. Successful matching across modalities is often considered to be a good indicator of vocabulary knowledge in children and measures the breadth of conceptual knowledge that a child has accumulated (Laws et al., 2014). The retrieval demands are relatively low, since the activation of the concept from its label should be relatively automatic. However, the presentation of the visual features of the object could generate conflict with the distractors and, therefore, the task could require controlled cognition to resolve selection demands and guide matching performance.

4.3.1 The experiential hypothesis and item matching

Although the probe items were selected to be familiar to children of all ages, the experiential hypothesis could predict a small advantage for matching in older children due to greater accumulated knowledge of the concepts. In the present findings, the performance of the oldest age group (96-130 months) approximated ceiling on this task, but the youngest age group (60-78 months) was less able to match across modalities than the middle age group (79-95 months). One arising issue is that the age advantage could not simply reflect more accumulated knowledge of concepts as the pattern was dissimilar to performance in the naming task that used identical items. On the naming task, the middle age group did not outperform the youngest age group but the oldest age group did.

4.3.2 The CSC framework and item matching

The CSC framework (Rogers et al., 2015) emphasises controlled cognition as supporting retrieval and selection on semantically-driven tasks. For item matching, the control demands of the task were relatively low compared to other tasks, with automatic activation of the visual features of known concepts after hearing the identifying word. However, by including distractors that increased conflict with the features of the target item, then controlled cognition was necessary in order to make an accurate selection. Older children should resist interference from distractors better than the youngest children, given their better capacity for invoking controlled cognition to support matching. In the present study, there was an age by distractor type interaction, where the older age group benefited when the distractors were proximally-related to the target, but not when distractors were distal. The younger 60-78-month-olds were therefore more susceptible to errors during item-matching, possibly because they were less able to recruit controlled cognition to overcome the featural similarity of semantically-related distractors.

A key prediction was that children's performance in relation to typical items would also vary by distractor type. More typical items should benefit from less feature overlap with distractors in the distal condition. Equally, atypical items should benefit from a lack of shared features with proximal distractors, compared to the likely feature overlap with distractors from other categories in the distal condition. The findings supported an interaction between typicality and distractor type. If the probe item was highly or mostly typical of its category, then children were more accurate in the distal condition, consistent with the prediction from the CSC framework. In other words, proximal distractors impeded younger children's ability to respond successfully to the more typical items due to greater competition from same-category distractors, consistent with the CSC framework (Rogers et al., 2015). However, for the atypical items, there was no advantage for matching under the proximal distractors or distal distractors. There was no evidence that younger children responded to atypical images in a way that created competition for selecting the correct match when distractors are selected from other more distal categories. This may be because atypical concepts bring to mind fewer items than typical ones (Rogers et al., 2015). Therefore, children were able to correctly match atypical probes, even in the presence of distal distractors.

4.4 Evaluating findings from the picture sorting task

Successful picture sorting requires the concept to be discriminated from concepts outside a cluster and grouped with concepts within a related cluster, to determine whether an object is a member of a selected category. The sortal term will determine the semantic relations that are most salient for the task. This will also point to the perceptual and semantic features that define category membership (Rosch et al., 1976). A sortal at the basic level involves more perceptual similarity and a greater number of shared semantic features between objects within a cluster (Rosch et al., 1976), compared to the semantic relations shared by a superordinate cluster. This implies that basic level sorting involves automatic activation of concepts within the cluster, whereas superordinate sorting involves access to broader semantic relations that could require more controlled retrieval. The specificity of a basic sortal term enables successful sorting for children aged 4-5 years (Mervis & Crisafi, 1982; Rosch et al., 1976) compared to superordinate sorting. By contrast, superordinate sorting is more vulnerable to task demands (Saxby & Anglin, 1983; Murphy & Brownell, 1985).

4.4.1 The experiential hypothesis and picture sorting

This view suggests that greater experience with objects allows more access to perceptual features in concept knowledge. Children were predicted to show benefits of age when sorting, with no strong apriori assumptions regarding preferences for the type of sortal. In the present study, the youngest age group (60-78 months) were more accurate when sorting using a specific sortal, compared to a general (superordinate) sortal term. According to the experiential view, the direct experience of the youngest children with identifying the perceptual features of items could allow more explicit identification of concepts and relative ease of discriminating basic level categories (see also Rosch et al.'s, 1976 oddity task). Alternatively, their early experience with thematic relations may have made the specific sort easier due to the categories used: air, land and water. For example, they may have learnt through observation that goldfish swim in water, associating goldfish with the water category. The older children were at ceiling for the general sortal, which could reflect their emerging familiarity with using taxonomic categories.

The pattern of typicality effects generated mixed support for the experiential view. A full interaction between age and typicality could not be determined due to ceiling effects of the older age groups on the general sortal. When using a specific sortal, typicality effects were identified that were consistent across age groups. Children were more accurate at sorting the most typical items, suggesting that early experience could provide clues to sorting these typical concepts. However, the effects of typicality were not graded; children also performed well when sorting *atypical* pictures. The absence of an interaction of typicality by age also ruled out the proposal from Jerger and Damian that younger children have less knowledge of atypical objects. Jerger and Damian (2005) only used two (typical and atypical) groupings to demonstrate a typicality advantage that could mirror the advantage observed here for the high typicality group as compared to the medium one. Since graded effects of typicality were not found, it seemed that the perceptual distinctiveness of typical and atypical items contributed to sorting into the specific, subordinate category level at all ages. Sorting by subordinate categories is therefore sensitive to the distinguishing features of the items in children.

4.4.2 The CSC framework and picture sorting

The CSC framework (Rogers et al., 2015) predicted a graded benefit for more typical items. According to the CSC framework (Rogers et al., 2015), low typicality items bring to mind items from other related clusters that invoke more control demands when attributing category membership. However, children in the present study were more able to attribute category membership to the most and least typical items. This pattern was unlike previous studies with SA and SD patients as Rogers and colleagues (2015) used three levels of typicality and obtained a graded effect of typicality. Neither the CSC framework (Rogers et al., 2015) nor the experiential hypothesis can fully account for this finding in children. In other ways, the findings were consistent with the CSC framework (Rogers et al., 2015), for example, when comparing a specific sortal to a general sortal term. The two older groups performed close to ceiling on the general sortal and were less accurate on the specific sortal. However, the youngest age group showed the opposite pattern; more errors were made on the general, compared to specific, sorting task. One explanation is that the youngest age group were challenged to provide controlled retrieval of more distant semantic relations between an item and its superordinate category. As younger children are less likely to implement controlled retrieval, then their generic sorting is more susceptible to error. Overall, the younger children were more susceptible to task demands due to less implementation of controlled retrieval, but, across age groups, children were generally more sensitive than adults to using distinctive characteristics for categorising items.

4.5 Evaluating findings from the category matching task

Category matching requires the induction of a 'rule' about category membership that allows the conjunction of a probe object to a target, in the face of competing alternatives on a single trial. In the current design, both proximal and distal conditions of the category matching task required the child to identify which category the probe item belonged to in order to make a successful match. However, these conditions varied by the semantic and featural relations shared with the distractors. In the distal condition, the defining features of the probe provide a strong clue to category membership without eliciting competition from related distractors. In the proximal condition, characterising features are more commonly shared with the distractors as members of a specific related category.

4.5.1 The experiential hypothesis and category matching

According to the experiential hypothesis, children should increase in their ability to identify taxonomic category relations as their knowledge of categories increases with age. Consistent with this view, the older age groups performed better than younger age groups at category matching. However, these graded age effects do not exclusively rule in improvements in knowledge of taxonomic relations, rather these age effects could also reflect differences in the child's ability to meet the contextual demands of the task.

Typicality could also increase children's ability to identify the taxonomic category relation, since more typical items would be more likely to be well-established into a semantic category at a younger age, compared to atypical items. Since typical and atypical items were matched for familiarity, it was likely that older children would be able to conduct category matching in a similar way. Only the younger children were predicted to struggle with matching atypical items, due to their developmental lag in co-ordinating category membership of atypical items. However, the age by typicality prediction was not evident: both older and younger children were able to co-ordinate knowledge of the typical compared to atypical items. Similar to picture naming, the typicality advantage could be explained in two ways. If knowledge of typical items is accumulated before that of atypical ones in a developmental progression, then the structure of the semantic network supports richer clusters for typical objects that are stable across childhood. Or, an enriched semantic network for highly typical objects, at the 'heart' of their cluster or category, generates more retrieval cues than less typical objects to facilitate category matching.

4.5.2 The CSC framework and category matching

The CSC framework (Rogers et al., 2015) predicted a graded benefit for more typical items in the distal condition since less typical items automatically bring to mind items from other related clusters, which invoke more control demands when attributing category membership (similar to category sorting). The CSC framework (Rogers et al., 2015) also predicted that children would be more accurate at category matching with distal, rather than proximal, distractors. The assumption is that competition between candidate distractors and target object is stronger in the proximal condition, due to the featural and semantic similarity of the distractors. Consistent with this view, probes were more likely to be matched successfully to their category associate, when distractors were distal rather than proximal. If distal distractors elicit lower levels of competition with typical category members, then a graded benefit for more typical items would occur as atypical items provide more opportunities for conflict with miscellaneous categories. By contrast, if control requirements increase further with proximal distractors that share featural and semantic relations with the target, then performance on more typical items will be depressed and the direction of the typicality advantage could favour better matching of *atypical* probe items due to reduced competition, predicting a typicality*condition interaction.

Interestingly, there was an interaction between distractor type and typicality, but it was not in the predicted direction. With proximal distractors, children showed a graded advantage for matching more typical items, but there was no typicality effect in the distal condition. Our finding relating to the proximal condition is surprising because distractors that share a close semantic relationship with a high typicality probe should implicate more, rather than less, competition according to the CSC framework. It is possible that rule induction of the category match was more favourable for children when the semantic and feature relations were primed by a visual display of similar distractors, therefore adding to the association strength between the highly typical probes and the targets and reducing the demands for controlled retrieval. The lack of a typicality effect in the distal condition suggests there were no differences in category matching, possibly because the items were well-matched in their relative familiarity. The absence of a graded effect of typicality of items when attributing category membership is similar to the findings observed for the category sorting task, where atypical items also showed a relative advantage compared to slightly more typical items. It is possible that children are more biased towards capturing the distinctiveness of the most and least typical items to facilitate category judgements, whether sorting or matching by category.

4.6 Evaluating findings from the object-use matching task

Matching by object use requires the induction of a 'rule' about the function of the probe object and its association with the recipient of that function. In the current design, both proximal and distal conditions of the object-use matching task required the child to identify the relevant probe item by its function to make a successful match to the recipient. In the distal condition, the defining features of the probe provide a strong clue to the rule as a sufficient basis for matching the object use. However, in the proximal condition, more specific knowledge of object use of the probe was required to discriminate from distractors with similar featural and semantic relations to the target. Since children aged four years can determine the function of an object (Deák et al., 2002), even the youngest age group would have attained knowledge of these object functions.

4.6.1 The experiential hypothesis and object-use matching

The experiential view predicted that all age groups would generate high levels of accuracy for the familiar items, given their likely exposure to the object functions, leading to a negligible effect of age. By contrast, the current findings favoured a graded advantage for older age groups matching more accurately than the younger age groups. It is possible that all children were aware of the object functions, but older children could infer the semantic relation to the recipient of the action more readily than younger children, suggesting a lack of experience of object recipients in younger children.

4.6.2 The CSC framework and object-use matching

According to the CSC framework (Rogers et al., 2015), the target objects should activate representations of other items that share similar features. This activation may be boosted when proximal distractors are presented. To select the correct recipient of the function of the probe object, controlled cognition is required to a greater extent in the proximal condition compared to the distal condition. Consistent with this view, children were more accurate with distal as compared to proximal distractors. The poorer scores on object-use matching by younger children could be explained by a difficulty with inferring the semantic relation of the recipient of the action, that implicates either a lack of knowledge of suitable recipients or weaker ability to invoke controlled cognition. In the absence of an interaction of age*distractor type, there was not sufficient evidence to rule-in an age-related shift in deploying controlled cognition for children more than 78 months on this task. Rather, there was a graded improvement in the performance for all children in older age groups.

4.7 Comparing findings from children and semantically-impaired adults

A key hypothesis of this study was that younger children would have reduced cognitive control to overcome intrinsic and extrinsic demands on retrieval of selected semantic tasks, compared to older children. In this way, the youngest children aged 5 years – 6 years 6 months were anticipated to rely more directly on automatic semantic processing from core knowledge and to reflect a similar pattern to adults with semantic impairment. Apart from the naming and object-use matching tasks, the findings broadly supported this view for three tasks; item-matching, category-matching and sorting. Specifically, for young children below 6 years 6 months, their performance on these semantically-driven tasks appeared to be linked to a significant difficulty with invoking controlled cognition, suggesting parallels with the profile of semantic aphasia in adults.

4.7.1 Comparing findings on the picture naming task

On the naming task, older children showed a positive advantage for naming the more typical objects, in contrast to the finding that adult controls are better at naming *atypical* objects (Rogers et al., 2015). Other studies with patients (Rossiter & Best, 2013) and healthy adults (Dell'Acqua et al., 2000) support the current finding. One possible source of the discrepancy is the categories of objects used. Rogers and colleagues (2015) focused on pictures of animals and vehicles; common general categories. These categories could bring to mind relatively more members of the same category, compared to other categories. This may explain a pattern of mixed findings for observing a typicality advantage across studies of patients (Rossiter and Best, 2013) and healthy adults (Dell'Acqua et al., 2000). However, the categories used by Rogers and colleagues (2015) were the same as those adopted in the present study with children.

One explanation for this age variation in the direction of the typicality effect is that the youngest children rely more on direct automatic access to feature cues for identification when providing object names, compared to older children and adults. More activation of feature cues supports retrieval when these objects are more typical of a category, rather than compromising retrieval for more typical items in older adults. For adults and possibly children older than those in the current sample, the likelihood of generating the appropriate object name becomes more likely for the atypical items since these features develop with age.

4.7.2 Comparing findings on the item matching task

On the item matching task, younger children showed a similar pattern of interaction between typicality of items and distractors on performance to that observed with SA and SD patients (Rogers et al., 2015). Younger children were more accurate with distal than proximal distractors, but only for more typical items. There was no impact of distractor type on the low typicality items. Rogers and colleagues (2015) similarly found that increasing typicality improved performance in the distal condition, but not in the proximal condition for SD patients and those with more severe SA. The overlapping profiles for these patients with more severe semantic impairment and all age groups of children, points to a common difficulty with identifying object identity under conditions of distraction, where the task invokes more requirements for controlled semantic retrieval. Both groups appeared to benefit from the reduced requirement for cognitive control in the distal condition. However, the typicality of the object was also important.

4.7.3 Comparing findings on the picture sorting task

On the picture sorting task, only the findings for the youngest children were comparable to the patients reported by Rogers and colleagues (2015). Older children were at ceiling on the general sorting task. The SA and SD patients were more accurate at sorting into specific categories if pictures were more, as compared to less, typical. Similarly, children were more accurate with high as compared to medium typicality images. However, there was also an advantage for the least typical items, compared to medium items, when sorting by a specific sub-category. When these children were asked to access specific feature knowledge to generate a successful sort to a specific category, a prevalence of distinguishing features of both atypical and typical items seemed beneficial. An interesting feature was that younger children were relatively disadvantaged by the general sortal that required access to more distant semantic relations between the object identity and the general category. Why is this? The general sortal requires access to taxonomic relations that are more 'distant' and lack a clear relation to distinguishing features (i.e. not all land animals have whiskers). Whereas the older children and patients were able to access these taxonomic relations to sort successfully, the youngest children were not able to do so, possibly because they were not able to invoke sufficient control to resolve the more distant semantic relations.

4.7.4 Comparing findings on the object-use matching task

For the object-use task, child performance was consistent with the pattern of object-use shown by SA patients who obtained low scores on semantic association tests (Corbett et al., 2009). Like other matching tasks, such as category matching, children were more accurate at matching images to their recipients when distractors were distal, rather than proximal, implying that children benefited when the task placed fewer demands on controlled cognition. Overall, it is likely that the object-use task placed maximum demand on controlled semantic cognition, compared to other matching tasks, due to the need to induce a distant semantic relation between the functional action of the object and its recipient. Corbett and colleagues (2009) found their SA patients were least successful when matching to object function, or the recipient of the action, compared to item-matching. In agreement with this finding, in a previous associative matching study (Jefferies & Lambon Ralph, 2006), matching accuracy was reduced for both SA and SD patients when the semantic relation was more difficult to determine. Overall, both children and semantically-impaired patients were challenged to use semantic information in a task-appropriate fashion, when the controlled retrieval demands of the task were high. However, there was no evidence that the youngest age group was markedly worse when the distractors were proximally related to targets.

4.8 Strengths, limitations and future directions

A key strength of this research is that it focuses on semantic control in children. Not many studies have addressed this issue so it makes a valuable contribution to the literature. Also, although typicality effects have been found in adult naming (Dell'Acqua et al., 2000), these effects have been neglected in naming studies involving children. Our experiment shows that children between the ages of five and ten, like adults, display a typicality advantage in naming. Despite the contribution our investigation makes, it leaves certain questions unanswered. For example, we do not know whether our effects would be replicated if other categories were used. Future research should investigate a broader range of categories. It should also aim to discover whether the effects that we found are present in older children.

4.9 Conclusions

This study contrasted two hypotheses regarding the origins of conceptual development across childhood; experiential and controlled semantic cognition (CSC) hypotheses. The study used identical objects across four tasks. Although this design has limitations, patterns of variation across tasks, or by age, could not be simply attributed to a lack of knowledge of the identity of objects used. Older children were noticeably at ceiling when sorting by general taxonomic relations, or when matching target words across modality to the item picture. The oldest age group also benefited from stronger picture naming, category and object matching, compared to younger children, which was either indicative of more established semantic knowledge, or being more able to resolve task demands using controlled cognition. For all tasks, apart from naming, the youngest age group (60-78 months) performed more poorly than a slightly older age group (79-95 months). Age effects were not consistently graded across tasks, suggesting that task demands rather than knowledge of the concepts were important.

Both hypotheses were compatible with conceptual knowledge as resembling a network of inter-related concepts that is characterised by a 'small-world' structure of inter-linked clusters. Typical concepts were assumed to be i) graded from atypical concepts and ii) differentiated by the density of their semantic relations and feature similarity to other concepts within a cluster or category. In the present study, graded typicality effects were found in picture naming, consistent with a denser cluster of features and relations for typical items, in children of all ages. There were mixed findings for a consistent graded typicality effect in other tasks. Younger children were more efficient at sorting the most and the least typical objects, possibly through the automatic activation of their defining features as clues to the object identity when engaging in a specific sorting task. However, for item matching and category matching, typicality was linked to the task context, implying a role for controlled cognition.

Only the CSC hypothesis involves the recruitment of controlled cognition to guide controlled retrieval and reduce selection demands when performance varies across different distractors. A key drawback of the experiential hypothesis is that it focusses on core knowledge of concepts and fails to acknowledge the influence of task demands. For example, if an individual knows what something is, they will be able to match the word for it to its picture regardless of distractor type. Consistent with the CSC hypothesis, all three matching tasks (by item, category and object-use) were susceptible to the task context of semantically-related distractors. When distractors are proximal, controlled processes are supposed to direct the retrieval of the object identity, and so to facilitate matching when the automatic processing of features or semantic relations provides multiple cues to matching. The present findings suggested that controlled processing influences accuracy in the youngest children, similar to the findings from SA and SD patients (Rogers et al., 2015). Stored knowledge of object identity alone was not sufficient for the successful completion of these matching tasks, rather the relevant semantic and featural knowledge must be activated in a task-appropriate fashion.

To conclude, the CSC framework (Rogers et al., 2015) makes a valuable contribution to theorising in the development of semantic cognition in childhood. The parallels between adult patients with semantic aphasia and young children aged 60-78 months suggest that controlled cognition is emerging around six and a half years of age. Below this age, children were more susceptible to weaker performance where the control requirements were higher. Future studies should inspect similar issues in a broader array of objects and categories and consider the ongoing development of semantic cognition in older children (above the age of 10). The emergence of a mature semantic control system is an important aspect of cognitive development and the strengths and limitations of this system, particularly for children's ongoing educational success, warrant further research.

5. References

Adams, C., Hesketh, A., Reeves, D., Cooke, R. E., & Crutchley, A. (2001). *Assessment of Comprehension and Expression (6-11).* (n.p.).

Adlington, R. L., Laws, K. R., & Gale, T. M. (2009). The Hatfield Image Test (HIT): A new picture test and norms for experimental and clinical use. Journal of Clinical and Experimental Neuropsychology, 31(6), 731-753. https://doi.org/10.1080/13803390802488103

Anderson, J. D. (2008). Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter. *Journal of Fluency Disorders, 33*(2), 135-155. https://doi.org/10.1016/j.jfludis.2008.04.001

Ashcraft, M. H. (1978). Property norms for typical and atypical items from 17 categories: A description and discussion. *Memory & Cognition, 6*(3), 227-232. https://doi.org/10.3758/BF03197450

Asher, Y. M., & Kemler Nelson, D. G. (2008). Was it designed to do that? children's focus on intended function in their conceptualization of artifacts. *Cognition, 106*(1), 474-483. https://doi.org/10.1016/j.cognition.2007.01.007

Assink, E. M. H., van Well, S., & Knuijt, P. P. N. A. (2003). Age-of-acquisition effects in native speakers and second-language learners. *Memory & Cognition, 31*(8), 1218-1228. https://doi.org/10.3758/BF03195805

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language, 20*(2), 395-418. https://doi.org/10.1017/S0305000900008345

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *The Quarterly Journal of Experimental Psychology, 50A*(3), 560-585. https://doi.org/10.1080/783663595

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. https://doi:10.18637/jss.v067.i01

Bjorklund, D. F., & Thompson, B. E. (1983). Category typicality effects in children's memory performance: Qualitative and quantitative differences in the processing of category information. *Journal of Experimental Child Psychology, 35*(2), 329-344. https://doi.org/10.1016/0022-0965(83)90086-3

Bjorklund, D. F., Thompson, B. E., & Ornstein, P. A. (1983). Developmental trends in children's typicality judgments. *Behavior Research Methods & Instrumentation, 15*(3), 350-356. https://doi.org/10.3758/BF03203657

Blaye, A., & Jacques, S. (2009). Categorical flexibility in preschoolers: Contributions of conceptual knowledge and executive control. *Developmental Science, 12*(6), 863-873. https://doi.org/10.1111/j.1467-7687.2009.00832.x

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal Of Psychology, 93*, 89-114. Retrieved from https://onlinelibrary.wiley.com/journal/20448295

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE, 5*(5), 1-13. https://doi.org/10.1371/journal.pone.0010773

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS ONE, 9*(9). https://doi.org/10.1371/journal.pone.0106953

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *The Quarterly Journal of Experimental Psychology, 64*(3), 545-559. https://doi.org/10.1080/17470218.2010.503374

Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition, 13*(7-8), 992-1011. https://doi.org/10.1080/13506280544000165

Corbett, F., Jefferies, E., Ehsan, S., & Lambon Ralph, M. A. (2009). Different impairments of semantic cognition in semantic dementia and semantic aphasia: Evidence from the non-verbal domain. *Brain, 132*(9), 2593-2608. https://doi.org/10.1093/brain/awp146

Corrigan, R., & Schommer, M. (1984). Form versus function revisited: The role of social input and memory factors. *Child Development, 55*(5), 1721-1726. https://doi.org/10.2307/1129919

Deák, G. O. (2000a). Hunting the fox of word learning: Why "constraints" fail to capture it. *Developmental Review, 20*(1), 29-80. https://doi.org/10.1006/drev.1999.0494

Deák, G. O. (2000b). The growth of flexible problem solving: Preschool children use changing verbal cues to infer multiple word meanings. *Journal of Cognition and Development, 1*(2), 157-191. https://doi.org/10.1207/S15327647JCD010202

Deák, G. O., Ray, S. D., & Pick, A. D. (2002). Matching and naming objects by shape or function: Age and context effects in preschool children. *Developmental Psychology, 38*(4), 503-518. https://doi.org/10.1037/0012-1649.38.4.503

Dell'acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers, 32*(4), 588-615. https://doi.org/10.3758/BF03200832

Denney, N. W., & Ziobrowski, M. (1972). Developmental changes in clustering criteria. *Journal of Experimental Child Psychology, 13*(2), 275-282. https://doi.org/10.1016/0022-0965(72)90090-2

Duncan, E. M., & Kellas, G. (1978). Developmental changes in the internal structure of semantic categories. *Journal of Experimental Child Psychology, 26*(2), 328-340. https://doi.org/10.1016/0022-0965(78)90012-7

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development, 65*(3), 903-917. https://doi.org/10.2307/1131427

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1103-1123. https://doi.org/10.1037/0278-7393.26.5.1103

Fenson, L., Vella, D., & Kennedy, M. (1989). Children's knowledge of thematic and taxonomic relations at two years of age. *Child Development, 60*(4), 911-919. https://doi.org/10.2307/1131032

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*(1), 116-124. Retrieved from https://link.springer.com/journal/13428

Fox, J., & Weisberg, S. (2011). *An R companion to applied regression, second edition.* Thousand Oaks CA: Sage. Retrieved from http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Freeman, K. E., & Sera, M. D. (1996). Reliance on visual and verbal information across ontological kinds: What do children know about animals and machines? *Cognitive Development, 11*(3), 315-341. https://doi.org/10.1016/S0885-2014(96)90007-6

Freund, L. S., Baker, L., & Sonnenschein, S. (1990). Developmental changes in strategic approaches to classification. *Journal of Experimental Child Psychology, 49*(3), 343-362. https://doi.org/10.1016/0022-0965(90)90064-F

Funnell, E., Hughes, D., & Woodcock, J. (2006). Age of acquisition for naming and knowing: A new hypothesis. *The Quarterly Journal of Experimental Psychology, 59*(2), 268-295. https://doi.org/10.1080/02724980443000674

Gardner, H. (1973). The contribution of operativity to naming capacity in aphasic patients. *Neuropsychologia, 11*(2), 213-220. https://doi.org/10.1016/0028-3932(73)90010-9

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*(1), 65-95. https://doi.org/10.1016/0010-0285(88)90025-4

Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language, 32*(2), 249-268. https://doi.org/10.1017/S0305000904006786

Gopnik, A., & Meltzoff, A. N. (1992). Categorization and naming: Basic-level sorting in eighteen-month-olds and its relation to language. *Child Development, 63*(5), 1091-1103. https://doi.org/10.2307/1131520

Goswami, U. (2008). *Cognitive development : The learning brain*. Hove: Psychology Press.

Hoffman, P. (2018). An individual differences approach to semantic cognition: Divergent effects of age on representation, retrieval and selection. *Scientific Reports, 8*(1), 1-13. https://doi.org/10.1038/s41598-018-26569-0

Hoffman, P., McClelland, J. L., & Lambon Ralph, M. A. (2018). Concepts, control, and context: A connectionist account of normal and disordered semantic cognition. *Psychological Review, 125*(3), 293-328. https://doi.org/10.1037/rev0000094

Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition, 13*(7-8), 884-910. https://doi.org/10.1080/13506280544000093

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain, 129*(8), 2132-2147. https://doi.org/10.1093/brain/awl153

Jerger, S., & Damian, M.F. (2005). What's in a name? Typicality and relatedness effects in children. *Journal of Experimental Child Psychology, 92*(1), 46-75. https://doi.org/10.1016/j.jecp.2005.04.001

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology, 68*, 601-625. https://doi.org/10.1146/annurev-psych-122414-033702

Kimball, S., & Mattis, P. (2012). GIMP 2.8 [Computer software]. Retrieved from https://download.gimp.org/mirror/pub/gimp/v2.8/windows/

Kuznetsova, A., Brockhoff, P.B., & Christensen, R. H. B. (2017) lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1-26. https://doi.org/10.18637/jss.v082.i13

Lambon Ralph, M. A. (2014). Neurocognitive insights on conceptual knowledge and its breakdown. *Philosophical Transactions Of The Royal Society B, 369*(1634), 1-11. http://dx.doi.org/10.1098/rstb.2012.0392

Laws, G., Briscoe, J., Ang, S. Y., Brown, H., Hermena, E., & Kapikian, A. (2014). Receptive vocabulary and semantic knowledge in children with SLI and children with Down syndrome. *Child Neuropsychology, 21*(4), 490-508. https://doi.org/10.1080/09297049.2014.917619

Mandler, J. M., & McDonough. L. (1993). Concept formation in infancy. *Cognitive Development, 8*(3), 291-318. https://doi.org/10.1016/S0885-2014(93)80003-C

Medin, D. L., Altom, M. W., & Murphy, T. D. (1984). Given versus induced category representations: Use of prototype and exemplar information in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(3), 333-352. Retrieved from http://psycnet.apa.org/PsycARTICLES/journal/xlm/44/7

Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. *Child Development, 53*(1), 258-66. https://doi.org/10.2307/1129660

Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 70-84. Retreived from http://psycnet.apa.org/PsycARTICLES/journal/xge/147/5

Nazzi, T., & Gopnik, A. (2001). Linguistic and cognitive abilities in infancy: When does language become a tool for categorization? *Cognition, 80*(3), B11-B20. https://doi.org/10.1016/S0010-0277(01)00112-3

Neary, K. R., Van de Vondervoort, J. W., & Friedman, O. (2012). Artifacts and natural kinds: Children's judgments about whether objects are owned. *Developmental Psychology, 48*(1), 149-158. https://doi.org/10.1037/a0025661

Nelson, K. (1974). Variations in children's concepts by age and category. *Child Development, 45*(3), 577-584. https://doi.org/10.2307/1127823

Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia, 33*(10), 1281-1303. https://doi.org/10.1016/0028-3932(95)00102-9

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39-57. Retrieved from http://psycnet.apa.org/PsycARTICLES/journal/xge/147/7

Office for Standards in Education, Children's Services and Skills. (2004). Inspection report warden hill primary school (Ofsted Publication No. 258303). Retrieved from https://reports.beta.ofsted.gov.uk/provider/21/115740

Office for Standards in Education, children's Services and Skills. (2014). School report churchdown village junior school (Ofsted Publication No. 443879). Retrieved from https://reports.beta.ofsted.gov.uk/provider/21/115512

Patterson, K., Kopelman, M. D., Woollams, A. M., Brownsett, S. L. E., Geranmayeh, F., & Wise, R. J. S. (2015). Semantic memory: Which side are you on? *Neuropsychologia, 76*, 182-191. https://doi.org/10.1016/j.neuropsychologia.2014.11.024

Posnansky, C.J., & Neumann, P. G. (1976). The abstraction of visual prototypes by children. *Journal of Experimental Child Psychology, 21*(3), 367-379. https://doi.org/10.1016/0022-0965(76)90067-9

Quinn, P. C., Westerlund, A., & Nelson, C. A. (2006). Neural markers of categorization in 6-month-old infants. *Psychological Science, 17*(1), 59-66. https://doi.org/10.1111/j.1467-9280.2005.01665.x

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/.

Raven, J. (2008). Coloured Progressive Matrices and Crichton Vocabulary Scale Manual. London: Pearson.

Rogers, T. T., & McClelland, J. L. (2008). Précis of semantic cognition: A parallel distributed processing approach. *Behavioral and Brain Sciences, 31*(6), 689-714. https://doi.org/10.1017/S0140525X0800589X

Rogers, T. T., Patterson, K., Jefferies, E., & Lambon Ralph, M. A. (2015). Disorders of representation and control in semantic cognition : Effects of familiarity, typicality, and specificity. *Neuropsychologia, 76*, 220-239. https://doi.org/10.1016/j.neuropsychologia.2015.04.015

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382-439. https://doi.org/10.1016/0010-0285(76)90013-X

Rossiter, C., & Best, W. (2013). "Penguins don't fly": An investigation into the effect of typicality on picture naming in people with aphasia. *Aphasiology, 27*(7), 784-798. https://doi.org/10.1080/02687038.2012.751579

Rostad, K., Yott, J., & Poulin-Dubois, D. (2012). Development of categorization in infancy: Advancing forward to the animate/inanimate level. *Infant Behavior & Development, 35*(3), 584-595. https://doi.org/10.1016/j.infbeh.2012.05.005

Saxby, L., & Anglin, J. M. (1983). Children's sorting of objects from categories of differing levels of generality. *The Journal of Genetic Psychology, 143*(1), 123-137. Retrieved from https://www.tandfonline.com/toc/vgnt20/current

Schnur, T.T., Schwartz, M. F., Brecher, A., & Hodgson, C. (2006). Semantic interference during blocked-cyclic naming: Evidence from aphasia. *Journal of Memory and Language, 54*(2), 199-227. https://doi.org/10.1016/j.jml.2005.10.002

Schröder, A., Gemballa, T., Ruppin, S., & Wartenburger, I. (2012). German norms for semantic typicality, age of acquisition, and concept familiarity. *Behavior Research Methods, 44*(2), 380-394. https://doi.org/10.3758/s13428-011-0164-y

Smiley, S. S., & Brown, A. L.. (1979). Conceptual preference for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology, 28*(2), 249-257. https://doi.org/10.1016/0022-0965(79)90087-0

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*(1), 13-19. https://doi.org/10.1111/1467-9280.00403

Snyder, H. R., & Munakata, Y. (2010). Becoming self-directed: Abstract representations support endogenous flexibility in children. *Cognition, 116*(2), 155-167. https://doi.org/10.1016/j.cognition.2010.04.007

Snyder, H.R., & Munakata, Y. (2013). So many options, so little control: Abstract representations can reduce selection demands to increase children's self-directed flexibility. *Journal of Experimental Child Psychology, 116*(3), 659-673. https://doi.org/10.1016/j.jecp.2013.07.010

Tunnicliffe, S. D., & Reiss, M. J. (1999). Building a model of the environment: How do children see animals? *Journal of Biological Education, 33*(3), 142-148. https://doi.org/10.1080/00219266.1999.9655654

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67*(6), 1176-1190. https://doi.org/10.1080/17470218.2013.850521

Wainwright, P. E., Leatherdale, S. T., & Dubin, J. A. (2007). Advantages of mixed effects models over traditional ANOVA models in developmental studies: A worked example in a mouse model of fetal alcohol syndrome. *Developmental Psychobiology, 49*(7), 664-674. **https://doi.org/10.1002/dev.20245**

Ware, E. A., & Booth, A. E. (2010). Form follows function: Learning about function helps children learn about shape. *Cognitive Development, 25*(2), 124-137. https://doi.org/10.1016/j.cogdev.2009.10.003

Waxman, S., & Gelman, R. (1986). Preschoolers' use of superordinate relations in classification and language. *Cognitive Development, 1(*2), 139-156. https://doi.org/10.1016/S0885-2014(86)80016-8

Williams, D. A. (1982). Extra-binomial variation in logistic linear models. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 31*(2), 144-148. https://doi.org/10.2307/2347977

Wilshire, C. E., & McCarthy, R. A. (2002). Evidence for a context-sensitive word retrieval disorder in a case of nonfluent aphasia. *Cognitive Neuropsychology, 19*(2), 165-186. https://doi.org/10.1080/02643290143000169

Wolfinger, R. (1993). Laplace's approximation for nonlinear mixed models. *Biometrika, 80*(4), 791-795. https://doi.org/10.2307/2336870

Wright Cassidy, K., & Kelly, M. H. (2001). Children’s use of phonology to infer grammatical class in vocabulary learning. *Psychonomic Bulletin & Review, 8*(3), 519-523. https://doi.org/10.3758/BF03196187

6. List of abbreviations

Abbreviation	Explanation
ANOVA	Analysis of variance
AoA	Age of acquisition
CCT	Camel and cactus test
CSC	Controlled semantic cognition
DCCS	Dimensional change card sort
GLMM	Generalised linear mixed-effects model
LMER	Linear mixed-effects regression
LOFTS	Levels of Familiarity, Typicality and Specificity
Ofsted	Office for Standards in Education
SA	Semantic aphasia
SD	Semantic dementia

Appendix A: The typicality subset of the LOFTS battery (items and information)

Typicality		
Low	Medium	High
Snail	Squirrel	Tiger
Hedgehog	Chimp	Mouse
Tortoise	Gorilla	Donkey
Frog	Elephant	Lion
Kangaroo	Hippopotamus	Cheetah
Bat	Camel	Deer
Crocodile	Badger	Leopard
Swan	Duck	Robin
Penguin	Pheasant	Kingfisher
Ostrich	Woodpecker	Magpie
Crab	Shark	Goldfish
Seahorse	Ray	Trout
Hovercraft	Oil tanker (sea)	Canoe
Submarine	Aircraft Carrier	Rowing Boat
Caravan	Tractor	Lorry
Parachute	Helicopter	Aeroplane

The typicality subset of the LOFTS battery includes animals and vehicles. It is comprised of 16 triplets of items. For each triplet, there is a single item with low typicality (i.e. a rating of 1-2 on a 7-point Likert scale), medium typicality (i.e. a rating of 2-4.5) and a highly typical item (i.e. a rating above 4.5). Within each triplet, all three items are matched for rated familiarity.

Appendix B: Answer sheet for the naming task

Item	Answers given by at least two adult participants
Hedgehog	Hedgehog
Chimpanzee	Chimpanzee
Trout	Trout
	Carp
	Salmon
Goldfish	Goldfish
Cheetah	Cheetah
	Leopard
Lorry	Lorry
	Van
	Truck
Rowing Boat	Rowing Boat
	Dinghy
	Canoe
Canoe	Canoe
	Kayak
Tortoise	Tortoise

	Turtle
Pheasant	Pheasant
	Turkey
	Chicken
Robin	Robin
Magpie	Magpie
	Sparrow
	Blackbird
Caravan	Caravan
	Trailer
	Car
Deer	Deer
	Stag
Kingfisher	Kingfisher
	Woodpecker
Hovercraft	Hovercraft
	Hoverboat
	Submarine
	Ship
Woodpecker	Woodpecker
Gorilla	Gorilla

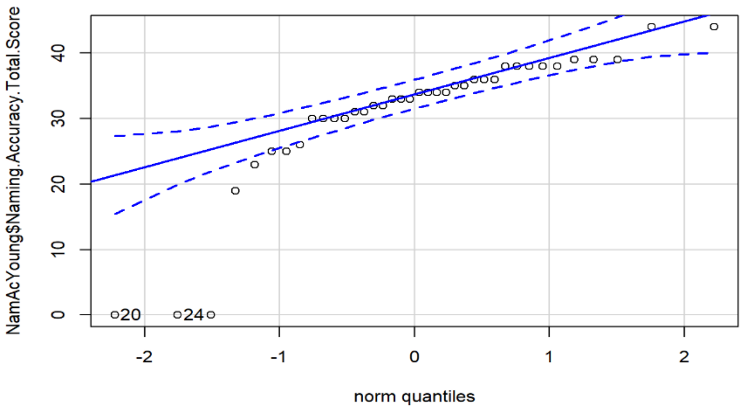
Aircraft Carrier	Aircraft Carrier
	Runway
	Ship
Leopard	Leopard
	Tiger
	Cheetah
Ostrich	Ostrich
	Emu
Swan	Swan
Lion	Lion
Helicopter	Helicopter
Tractor	Tractor
Tiger	Tiger
Mouse	Mouse
	Gerbil
Crocodile	Crocodile
	Alligator
Frog	Frog
	Toad
Shark	Shark
Hippo	Hippo

	Rhinoceros
Parachute	Parachute
	Paraglider
Submarine	Submarine
Oil Tanker	Oil Tanker
	Ship
	Cargo Ship
	Cruise Ship
	Battleship
Badger	Badger
	Skunk
Donkey	Donkey
Ray	Ray
Duck	Duck
Aeroplane	Aeroplane
Penguin	Penguin
Squirrel	Squirrel
Kangaroo	Kangaroo
Bat	Bat
Elephant	Elephant
Camel	Camel

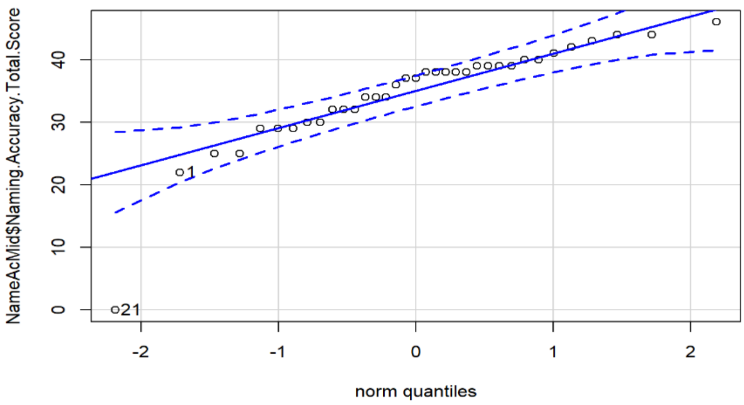
Snail Seahorse Crab	Snail Seahorse Crab

Appendix C: Q-Q plots for the naming accuracy data, split by age group

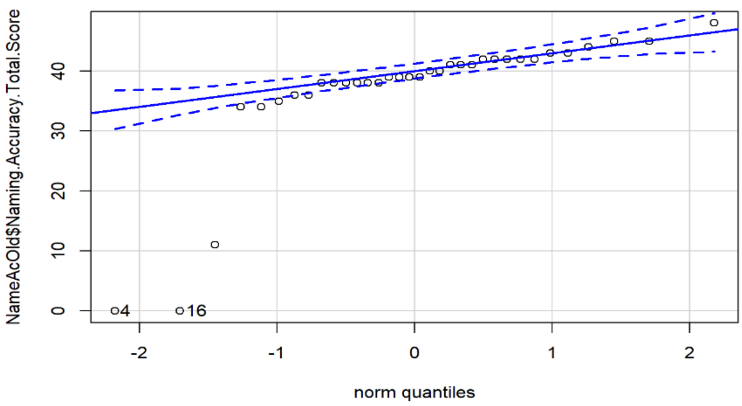
Group 1 (60-78 months)



Group 2 (79-95 months)

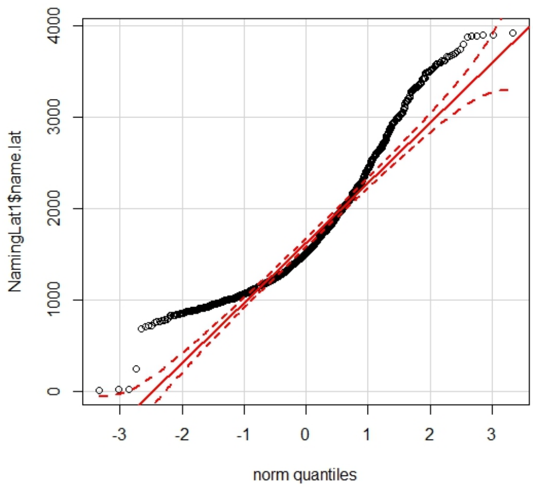


Group 3 (96-130 months)

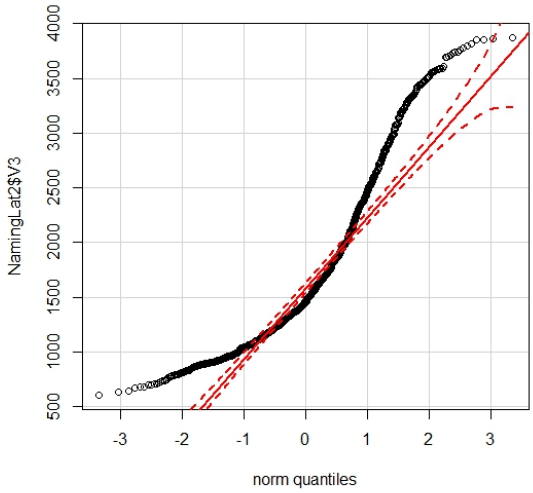


Appendix D: Q-Q plots for the naming latency data, split by age group.

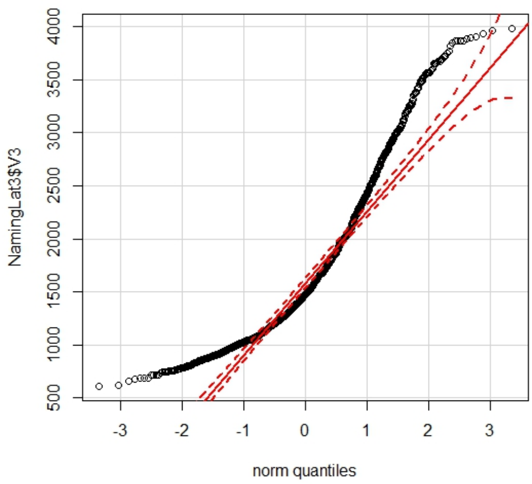
Group 1 (60-78 months)



Group 2 (79-95 months)

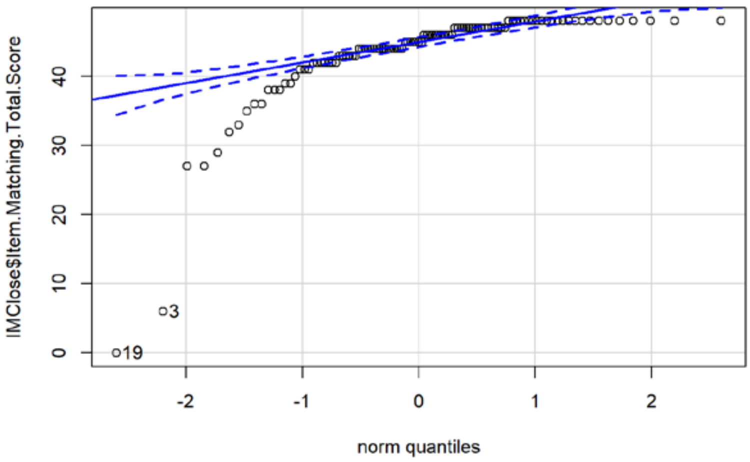


Group 3 (96-130 months)

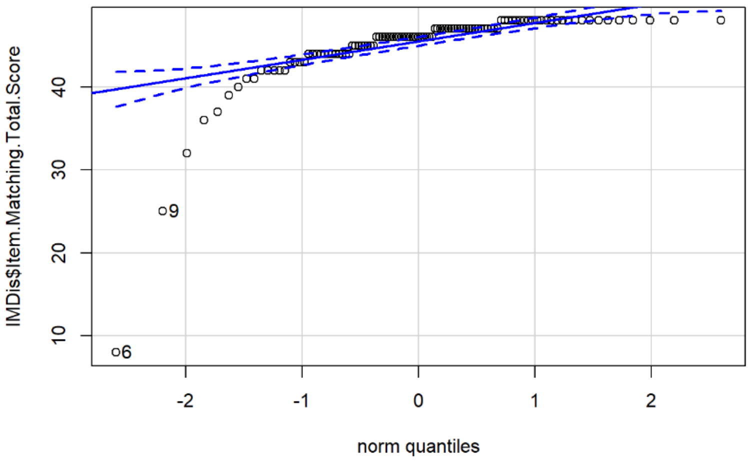


Appendix E: Q-Q plots for the item (word-picture) matching data, split by distractor type

Proximal distractors

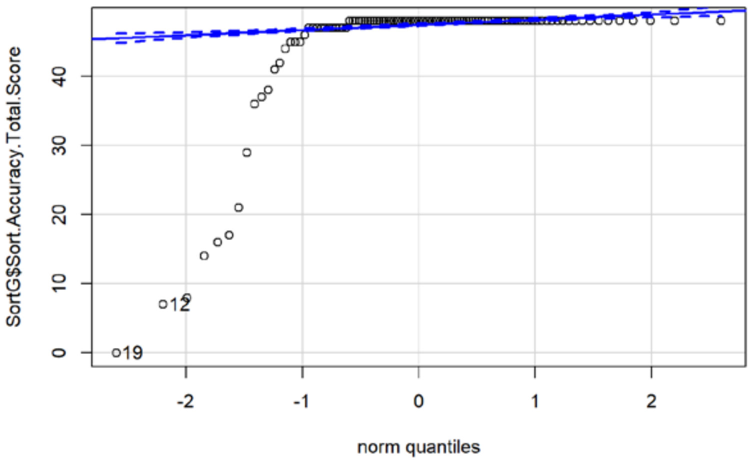


Distal distractors

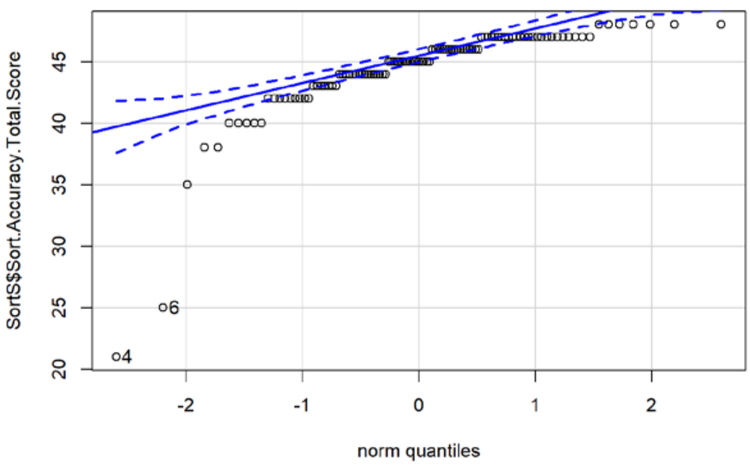


Appendix F: Q-Q plots for the sorting data, split by sortal type

General sortal

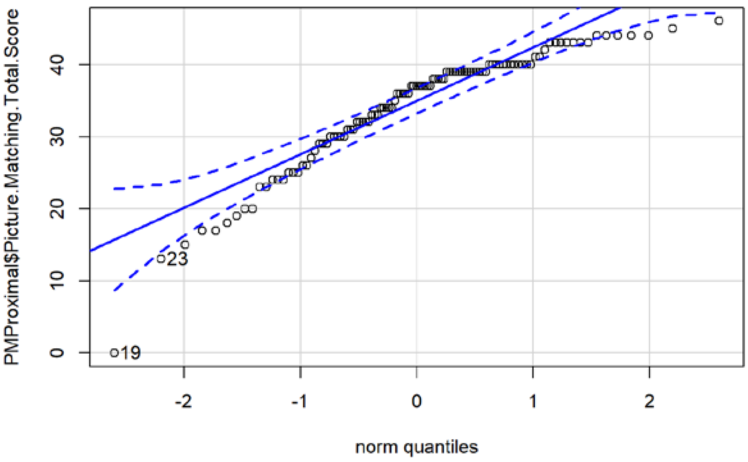


Specific sortal

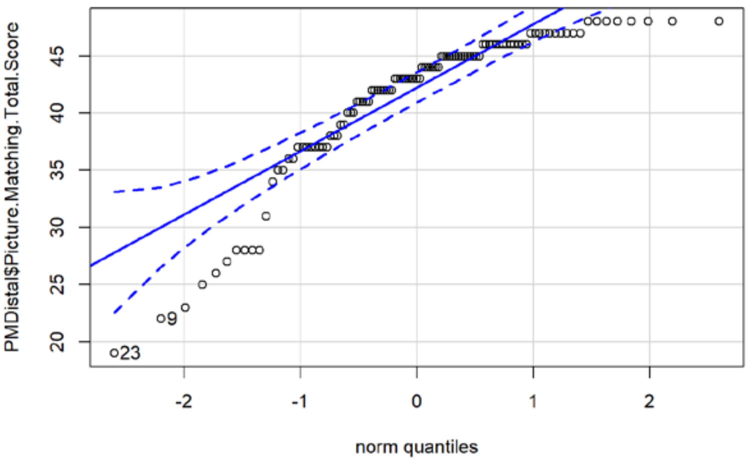


Appendix G: Q-Q plots for the category matching data, split by distractor type

Proximal distractors

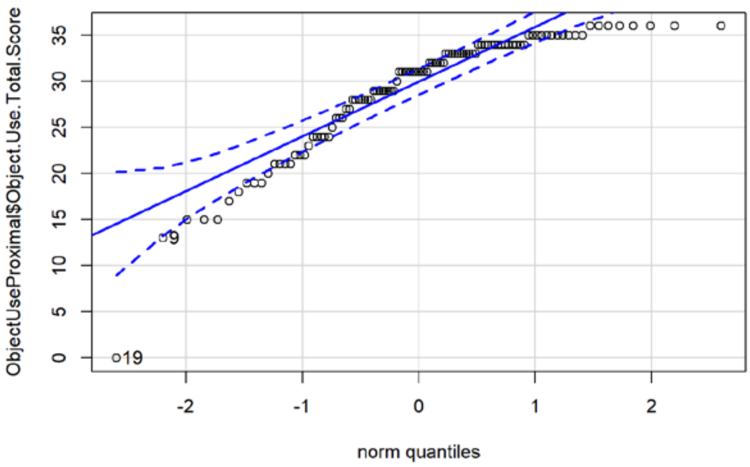


Distal distractors



Appendix H: Q-Q plots for the object-use data, split by distractor type

Proximal distractors



Distal distractors